SERC-THEP 2012 (Prep)

Particle Physics and the Standard Model

Anirban Kundu
Department of Physics, University of Calcutta,
92, Acharya Prafulla Chandra Road, Kolkata 700 009, India.
E-mail: akphy@caluniv.ac.in

Jyoti Prasad Saha
Department of Physics, Kalyani University

Contents

1 Dirac Equation **4**

1.1 More identities with γ matrices 5

1.2 Chirality and γ_5 6

1.3 Dirac equation and the gauge transformation 7

1.4 Free-particle solutions 8

2 Discrete Symmetries **8**

2.1 Parity 8

2.2 Time reversal 9

2.3 Charge conjugation 10

2.4 Symmetries in field theory 10

2.5 Spin-1 fields 11

2.6 Spin-$\frac{1}{2}$ fields: parity 12

2.7 Spin-$\frac{1}{2}$ fields: charge conjugation 13

2.8 Spin-$\frac{1}{2}$ fields: time reversal 14

2.9 Bilinear covariants 14

3 SU(3) and the Quark Model **16**

3.1 SU(3) 16

3.2 The quark model 17

3.3 Strong isospin 18

3.4 The SU(3) of flavour 19

3.5 On the mass of the quarks 22

3.6 P and C of mesons 23

3.7 Meson wavefunctions 23

3.8 Meson decays and the Okubo-Zweig-Iizuka rule 24

3.9 Baryons 25

3.10 Magnetic moment of baryons 26

3.11 Colour 27

4 Weak Interaction **27**

4.1 Fermi's theory of allowed β-decay: nonrelativistic treatment 28

4.2 Wu's experiment 32

4.3 Goldhaber experiment 33

4.4 Lessons from the Goldhaber experiment 34
4.5 Digression: kinematics of scattering and decay 35
    4.5.1 $2 \to 2$ scattering 36
    4.5.2 Mandelstam variables 36
    4.5.3 Cross-section 37
    4.5.4 Decay width 39
    4.5.5 Collider experiments 39
4.6 π^- decay 42
4.7 Muon decay 43

5 The Electroweak Theory **43**
5.1 Spontaneous breaking of a discrete symmetry 44
5.2 Spontaneous breaking of a continuous symmetry: Goldstone theorem 46
5.3 Spontaneous breaking of an abelian gauge symmetry: Higgs mechanism 47
5.4 Spontaneous breaking of a non-abelian gauge symmetry 49
5.5 The electroweak unification 49
5.6 The mass of the weak gauge bosons 51
5.7 Recovering the photon 52
5.8 Fermion masses 53

6 The Higgs Boson **54**
6.1 Theoretical bounds on the Higgs boson mass 54
    6.1.1 Unitarity of gauge boson scattering 54
    6.1.2 Triviality 56
    6.1.3 Vacuum stability 57
6.2 Production of the Higgs boson 57
6.3 Decay of the Higgs boson 60
6.4 Problems with the Higgs boson 60
6.5 The Standard Model 62

7 Flavour Physics **63**
7.1 The Cabibbo mechanism 64
7.2 Flavour-changing neutral current 65
7.3 The Glashow-Iliopoulos-Maiani (GIM) mechanism 65
7.4 The K meson system 67

7.5 Digression: two-level quantum mechanics 68
7.6 CP violation in the K system 70
7.7 The Cabibbo-Kobayashi-Maskawa matrix 71
7.8 The Wolfenstein parametrisation 73
7.9 $B^0 - \overline{B}^0$ mixing: box diagram 75
7.10 CP violation in the B system 79

8 Neutrinos **80**
8.1 The solar neutrino problem 80
8.2 The atmospheric neutrino problem 82
8.3 Neutrino mixing and hierarchy 82
8.4 The see-saw mechanism 83

9 The Structure of Proton **84**
9.1 Electron-muon scattering in the lab frame 84
9.2 Inelastic *ep* scattering 86

10 The Parton Model **87**
10.1 Parton kinematics 88
10.2 Parton distribution functions 90

Before we start

Summary of notations and conventions

- We use the natural system of units: $\hbar = c = 1$, and also $\varepsilon_0 = \mu_0 = 1$. Thus, the fine structure constant $\alpha = e^2/4\pi$. In this system, $[M] = [L]^{-1} = [T]^{-1}$, action is dimensionless, and Lagrangian density has dimension M^4.
- The flat space-time metric is $\eta_{\mu\nu} = \text{diag}\ (1,-1,-1,-1)$ and $ds^2 = dt^2 - dx^2 - dy^2 - dz^2$. This is also known as the Bjorken-Drell metric. $A_\mu = \eta_{\mu\nu}A^\nu$. Also, $\eta^{\mu\alpha}\eta_{\mu\nu} = \delta^\alpha_\nu$, so $\eta^{\mu\nu} = \text{diag}\ (1,-1,-1,-1)$ too.
- All Greek indices run from 0 to 3, Latin indices from 1 to 3 (space part only). Repeated indices are implicitly summed over.

There are some very useful webpages that the reader should be familiar with if (s)he wants to be a particle physicist. The first is the Particle Data Group website `http://pdg.lbl.gov`; this lists the properties of all known particles, elementary or composite, provides bounds on the hypothetical or yet-to-be-discovered ones, and contains a number of up-to-date reviews. Then comes the electronic archive `http://arXiv.org` (X is a lookalike for the greek 'chi' in archive). This is an excellent repository of physics papers; nowadays, almost everyone puts his or her papers in the arXiv, before submitting them to the journal. There are various categories; for our purpose, hep-ph, hep-ex, and hep-lat will suffice. It is free and so you can download at will (but beware, pre-arXiv papers are generally not available). The last one is INSPIRE `http://inspirehep.net`, which is a standard search engine for high-energy and astrophysics papers, and a very efficient one too. There are also a number of popular websites; they may be found from the Particle Data Group page. You can also google.

(There will be a major deviation from any standard particle physics textbook; after July 4, 2012, we are in the post-Higgs era, the Higgs boson has been as good as discovered and the theory known as the Standard Model is now complete.)

Prerequisites

I will assume that you know the following things, either from your M.Sc. days, or from the courses on QFT and Group Thoery. If you don't, it will be very difficult, if not outright impossible, for you to follow the Particle Physics course.

1. Energy-momentum relationship of special relativity, and the four-vector notation.
2. Basic ideas of relativistic electrodynamics, like gauge invariance, and field strength tensor $F^{\mu\nu}$.
3. Derivation of the Euler-Lagrange equations from the Lagrangian for point particles and Lagrangian density for classical fields.
4. Relativistic quantum mechanics, Klein-Gordon and Dirac equations. Algebra of the Dirac γ-matrices. Properties of the Dirac spinor, like spin-$\frac{1}{2}$ nature.
5. Canonical quantisation of spin-0 and spin-$\frac{1}{2}$ fields. For the basic ideas, see below. Also, the procedure to write down the Feynman rules.

6. Amplitude calculations for the basic QED processes, like $e\mu \to e\mu$, Moeller, and Bhabha scattering.
7. Idea about simplest Lie groups — SU(2) will do.

Essence of canonical quantisation

- Take the wave equation for a particle and re-interpret it as a wave equation for a field whose excitations are to be identified with the particle.
- Write the Lagrangian density, and check that the Euler-Lagrange prescription yields the correct equation of motion.
- Find the canonically conjugate momentum $\pi = \partial\mathcal{L}/\partial\dot{\phi}$ for each field ϕ.
- Start from the commutation relation between ϕ and π, which is a generalization of $[x_i, p_j] = i\hbar\delta_{ij}$.
- Make a Fourier expansion of the field in terms of plane waves. This is equivalent to assuming the field as an assemblage of infinite number of harmonic oscillators. They form a basis. The expansion looks like
$$\phi(x) \to C_1(\mathbf{p})a(\mathbf{p})e^{-ipx} + C_2(\mathbf{p})b^\dagger(\mathbf{p})e^{ipx}$$
with a suitable integration over $d^3\mathbf{p}$. $C_1 = C_2 = 1$ for spin-0 field; $C_1 = u(\mathbf{p})$, $C_2 = v(\mathbf{p})$ for spin-1/2 field; and $C_1 = \varepsilon_\mu(\mathbf{p})$, $C_2 = \varepsilon^*_\mu(\mathbf{p})$ for photon field.
- For a self-conjugate field, $a = b$, and the particle is its own antiparticle. Otherwise $a \neq b$. $a^\dagger$ creates a particle, a annihilates it; $b^\dagger$ creates an antiparticle, b annihilates it.
- From the $[\phi, \pi]$ commutator, find $[a, a^\dagger]$.
- For spin-1/2 fields, $\{\phi, \pi\}$ and $\{a, a^\dagger\}$ anticommutators are to be considered. This is consistent with the exclusion principle.
- $a^\dagger a$ is the number operator. To have a finite zero-point energy of the system, use the normal-ordered product, denoted by $:\phi(x_1)\phi(x_2):$.
- For quantisation of electromagnetic field, we have to choose a gauge to remove two spurious degrees of freedom in A_μ if it is to represent a real photon. This needs explicit breaking of manifest Lorentz invariance, but the invariance is restored in the final result. One might also do a gauge-invariant quantisation, but that is more complicated.
- The propagator is formally defined as the vacuum expectation value of the time-ordered product of two fields: $\langle 0|T(\phi(x_1)\phi(x_2))|0\rangle$. If ϕ is not self-conjugate, use ϕ^*, or $\bar{\psi}$ for a fermion.
- A shorthand prescription to get the propagator is to take the inverse of the wave equation in momentum space and multiply it by i. An extra term $+i\varepsilon$ is to be added in the denominator to shift the pole from the axis. Ultimately we take $\varepsilon \to 0$.
- The photon propagator is gauge dependent. Generally one uses the Feynman gauge: $iD_F^{\mu\nu}(p) = -i\eta^{\mu\nu}/p^2$. The final results must be independent of the gauge choice.

- This shows another way of writing the propagator in momentum space: $i\Sigma/p^2 - m^2$ where Σ is the sum of all spin states — 1 for spin-0 field, $\sum u\bar{u} = p_\mu\gamma^\mu + m$ for spin-1/2 field, and $\sum \varepsilon_\mu\varepsilon_\nu^* = -\eta_{\mu\nu}$ for photon field (in Feynman gauge).

Feynman rules for amplitude calculation

- Each Feynman diagram has three parts: external legs, vertices and internal propagators.
- The factor for scalar external leg is 1. For vector leg (massive or massless) it is ε_μ (for incoming) and ε_μ^* (or outgoing). For incoming spin-1/2 particle it is u ($\bar{v}$ if it is an antiparticle) and for outgoing particle it is $\bar{u}$ (v for an antiparticle).
- The factor associated with each vertex is just i times the term in the Lagrangian density, except the field operators.
- Internal propagators are $i/(p^2 - m^2)$ for scalar, $i/(\not{p} - m)$ for fermion, and $-i\eta_{\mu\nu}/p^2$ for photon in Feynman gauge. For a massive particle it is $-i\eta_{\mu\nu}/(p^2 - m^2)$.
- For a fermion current start from where a particle leaves the system and go against the fermion arrow. Put u, $\bar{u}$ and vertex factors in proper order. The factors associated with bosonic propagators can be put anywhere.
- Four-momentum conservation holds at each vertex. However, for any momentum l not constrained by energy-momentum relationship (i.e., for a loop) attach a factor of $\int d^4l/(2\pi)^4$.
- For a fermion loop put an extra factor of -1 by hand.
- Multiply the amplitude by the proper symmetry factor.
- While computing the amplitude squared, for unpolarized beams average over the initial spin states. Always sum over the final spin states unless you wish to project a particular spin state. Use completeness relations $\sum u\bar{u} = \not{p} + m$ and $\sum \varepsilon_\mu\varepsilon_\nu^* = -\eta_{\mu\nu}$.
- Summing over the spins means you can use the completeness relations (see text) and can express the square of any fermionic current as a trace of some γ-matrices. The remaining dot products involving the four-momenta can be expressed in terms of the Mandelstam variables (again, see text).

1 Dirac Equation

This is not exactly a part of the course but some sort of a prerequisite. You already know that the Dirac equation (DE)

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \tag{1}$$

can be obtained as the equation of motion from the Dirac Lagrangian

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi\,, \tag{2}$$

where $\bar{\psi} \equiv \psi^\dagger \gamma^0$. One can also have the equation w.r.t. $\bar{\psi}$, the adjoint equation,

$$\bar{\psi}\,(i\partial_\mu \gamma^\mu + m) = 0\,, \tag{3}$$

where the derivative acts on the left. This is a free-particle equation and one can write down the plane-wave solutions [1]. The spinor ψ is a four-component object. For massless fermions, we can do only with two-component spinors. Such objects are known as *Weyl spinors.* They are nothing but the 2-dimensional representations of the Lorentz group: $(\frac{1}{2}, 0)$ or $(0, \frac{1}{2})$, and are not invariant under parity. So, technically speaking, even if the electron were massless, we had to work with the 4-component Dirac spinor, the real $(1, 1)$ representation of the Lorentz group, because electromagnetism respects parity.

The DE can be written with more brevity:

$$(i\not{\partial} - m)\psi = 0\,, \tag{4}$$

where the Feynman slash is a shorthand notation for contraction with γ^μ:

$$\not{A} = A_\mu \gamma^\mu = A^\mu \gamma_\mu. \tag{5}$$

The DE is actually a set of four differential equations, coupling four components of ψ:

$$\sum_{k=1}^{4} \Big[\sum_\mu i(\gamma^\mu)_{jk}\partial_\mu - m\delta_{jk}\Big]\psi_k = 0. \tag{6}$$

The algebra is

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\,. \tag{7}$$

We further have

$$\gamma^{0\dagger} = \gamma^0, \quad (\gamma^0)^2 = 1\,, \quad \gamma^{i\dagger} = -\gamma^i = \gamma_i\,, \quad (\gamma^i)^2 = -1. \tag{8}$$

Since γ^0 anticommutes with three spatial γ^is, we have

$$\gamma^0 \gamma^\mu \gamma^0 = \gamma^{\mu\dagger}. \tag{9}$$

At this point we would also introduce the matrix γ_5 defined as

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \frac{i}{4!}\epsilon_{\alpha\beta\mu\nu}\gamma^\alpha\gamma^\beta\gamma^\mu\gamma^\nu\,, \tag{10}$$

---

[1] DE is not a single-particle equation. Whenever you put special relativity together with quantum mechanics, what you get is quantum field theory. The correct way to proceed is to deal with a Dirac field, with possibilities of particle and antiparticle creation and annihilation.

where ϵ_{0123} and its even permutations are $+1$, odd permutations are -1, and any two identical indices make it zero. Also, $\epsilon_{0123} = -\epsilon^{0123}$. γ_5 is hermitian, as can be checked explicitly.

The matrix γ_5 anticommutes with all other γ-matrices:

$$\{\gamma^\mu, \gamma_5\} = 0\,, \tag{11}$$

and $(\gamma_5)^2 = \mathbf{1}$, the 4×4 unit matrix.

Q. Prove the following identities:

1. $Tr(\gamma^\mu) = 0\,,$
2. $Tr(\text{odd numbers of } \gamma\text{s}) = 0\,,$
3. $Tr(\gamma^\mu\gamma^\nu) = 4\eta^{\mu\nu}\,,$
4. $\gamma_\mu\gamma^\mu = 4\,,$
5. $Tr(\gamma^\alpha\gamma^\beta\gamma^\mu\gamma^\nu) = 4[\eta^{\alpha\beta}\eta^{\mu\nu} - \eta^{\alpha\mu}\eta^{\beta\nu} + \eta^{\alpha\nu}\eta^{\beta\mu}]\,,$
6. $\gamma_\mu\gamma^\alpha\gamma^\mu = -2\gamma^\alpha\,,$
7. $\gamma_\mu\gamma^\alpha\gamma^\beta\gamma^\mu = 4\eta^{\alpha\beta}\,,$
8. $\gamma_\mu\gamma^\alpha\gamma^\beta\gamma^\nu\gamma^\mu = -2\gamma^\nu\gamma^\beta\gamma^\alpha\,,$
9. $Tr(\gamma_5) = 0\,,$ [Again, write γ_5 as the product of four γ-matrices and push γ^0 to the right-hand side.]
10. $Tr(\gamma^\mu\gamma^\nu\gamma_5) = 0\,.$ [The proof goes as follows. Suppose $\mu \neq \nu$. Then the product, with the ambiguity of a sign, is nothing but that of two γ matrices with different index, and that is zero. If $\mu = \nu$, it is just the trace of γ_5 — again with the sign ambiguity — but that is zero.]

1.1 More identities with γ matrices

Let us collect here a number of important relations involving the γ matrices, apart from those just given above. We define the combination

$$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]\,. \tag{12}$$

The following relations can be proved (try!)

$$\begin{aligned}
\sigma^{\mu\nu}\gamma_5 &= -\frac{i}{2}\epsilon^{\mu\nu\alpha\beta}\sigma_{\alpha\beta}\,,\\
\gamma^\mu\gamma^\nu\gamma^\alpha &= \eta^{\mu\nu}\gamma^\alpha - \eta^{\mu\alpha}\gamma^\nu + \eta^{\nu\alpha}\gamma^\mu + i\epsilon^{\mu\nu\alpha\beta}\gamma_5\gamma_\beta\,,\\
\sigma^{\mu\nu}\gamma^\alpha &= i\left(\eta^{\nu\alpha}\gamma^\mu - \eta^{\mu\alpha}\gamma^\nu + i\epsilon^{\mu\nu\alpha\beta}\gamma_5\gamma_\beta\right)\,.
\end{aligned} \tag{13}$$

Multiply the last equation by γ^δ from right, multiply the expression of $\sigma^{\mu\nu}\gamma^\delta$ by γ^α, and subtract, to get

$$\begin{aligned}
-2i\sigma^{\mu\nu}\sigma^{\alpha\delta} &= i(2\eta^{\nu\alpha}\eta^{\mu\delta} - 2\eta^{\mu\alpha}\eta^{\nu\delta} - i\eta^{\nu\alpha}\sigma^{\mu\delta} - i\eta^{\mu\delta}\sigma^{\nu\alpha}\\
&\quad + i\eta^{\mu\alpha}\sigma^{\nu\delta} + i\eta^{\nu\delta}\sigma^{\mu\alpha} + 2i\epsilon^{\mu\nu\alpha\delta}\gamma_5\\
&\quad - \frac{i}{2}\epsilon^{\mu\nu\alpha\lambda}\epsilon^{\beta\delta\rho\tau}\eta_{\lambda\beta}\sigma_{\rho\tau} + \frac{i}{2}\epsilon^{\mu\nu\delta\lambda}\epsilon^{\beta\alpha\rho\tau}\eta_{\lambda\beta}\sigma_{\rho\tau})\,.
\end{aligned} \tag{14}$$

1.2 Chirality and γ_5

In the Weyl (chiral) representation, the matrix γ_5 is diagonal:

$$\gamma_5 = \begin{pmatrix} -\mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} . \tag{15}$$

Thus no other γ-matrix can be diagonal, as γ_5 anticommutes with all other γ^μs. We define two operators

$$P_R = \frac{1}{2}(1+\gamma_5) = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} , \quad P_L = \frac{1}{2}(1-\gamma_5) = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} . \tag{16}$$

With $(\gamma_5)^2 = \mathbf{1}$, it is easy to see that P_R and P_L have the necessary properties of orthogonal projection operators:

$$P_R^2 = P_R\,, \quad P_L^2 = P_L\,, \quad P_R P_L = P_L P_R = 0\,, \quad P_R^2 = P_R\,, \quad P_L^2 = P_L\,. \tag{17}$$

What do they project out? Obviously, P_L (P_R) projects out the upper (lower) two components of the four-component Dirac spinor. Writing the four-component spinor ψ as the direct sum of two two-component Weyl spinors (which we can always do because the $(1,1)$ representation of the Lorentz group is nothing but a direct sum of $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$), we get

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} , \quad P_L\psi = \begin{pmatrix} \psi_L \\ 0 \end{pmatrix} = \psi_L\,, \quad P_R\psi = \begin{pmatrix} 0 \\ \psi_R \end{pmatrix} = \psi_R\,. \tag{18}$$

Any four-component Dirac spinor can thus be written as the sum of two Weyl spinors: $\psi = \psi_R + \psi_L$. These projections are called *chiral* projections [2].

The interpretation of the four components of the Dirac spinor in the Dirac-Pauli representation follows the standard particle-antiparticle idea: the upper two components conventionally correspond to the particle degrees of freedom (a spin-$\frac{1}{2}$ particle has two degrees of freedom) and the lower two to the antiparticle degrees. These two representations, Dirac-Pauli and Weyl, are related by a similarity transformation, which also acts on the spinor. Thus, it is incorrect to say that ψ_L is the particle and ψ_R is the antiparticle!

Q. Write $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ and show that $\gamma_5^\dagger = \gamma_5$. Hence show that $\psi_L^\dagger = \psi^\dagger P_L$ but $\psi_L^\dagger\gamma^0 = \psi^\dagger\gamma^0 P_R$.

Q. Show that

$$\begin{aligned} \bar\psi\psi &= \bar\psi_R\psi_L + \bar\psi_L\psi_R\,, \\ \bar\psi\gamma^\mu\psi &= \bar\psi_R\gamma^\mu\psi_R + \bar\psi_L\gamma^\mu\psi_L\,. \end{aligned} \tag{19}$$

Hence argue that if under some group ψ_R is a singlet but ψ_L is not, the kinetic term of the Dirac Lagrangian is allowed but the mass term is not.

Q. Show that $\mathcal{L}_{Dirac}$ is not hermitian. But also show that the difference from its hermitian conjugate is a four-derivative. Since this vanishes when we integrate over d^4x to get the action, $\mathcal{L}_{Dirac}$ is as good as a hermitian form.

Q. Suppose you start with a Lagrangian $\mathcal{L} = \bar\psi(i\partial_\mu\gamma^\mu + m)\psi$, *i.e.*, a Lagrangian with a wrong-sign mass term. Show that by the chiral rotation defined as $\psi' = \gamma_5\psi$, one can cast it in the form of the Dirac Lagrangian written in terms of ψ'. Use the properties of γ_5. *As ψ' is as good a spinor*

---

[2] From Greek “chiros”, meaning *hand*, which is why they are also called right- and left-handed spinors.

as ψ, this shows that the sign of the mass term in the Dirac Lagrangian is not important; it can always be fixed with a chiral rotation. This is not the case for a spin-0 particle, and a wrong-sign mass term has deeper implications.

Q. Show that the hermitian conjugates of $\bar\psi_1\gamma^\mu(1-\gamma_5)\psi_2$ and $\bar\psi_1(1-\gamma_5)\psi_2$ are $\bar\psi_2\gamma^\mu(1-\gamma_5)\psi_1$ and $\bar\psi_2(1+\gamma_5)\psi_1$ respectively. [This is expected. In the first case, both ψ_1 and ψ_2 are left-chiral; in the second case, ψ_1 is right-chiral and ψ_2 is left-chiral. Hermitian conjugation cannot change the chirality of a fermion.]

1.3 Dirac equation and the gauge transformation

The Dirac Lagrangian is obviously invariant under a global phase transformation $\psi(x) \to \exp[i\alpha]\psi(x)$ where α is a constant. However, if the transformation is local, *i.e.*, α is a function of space-time coordinates $\alpha = \alpha(x)$, the transformation does not keep the Lagrangian invariant. There is an extra term

$$\mathcal{L} \to \mathcal{L}' = \mathcal{L} - (\partial_\mu\alpha(x))\bar\psi\gamma^\mu\psi\,. \tag{20}$$

The only way to cancel this extra term is to consider the particle in an electromagnetic field $A_\mu(x)$. From classical electrodynamics, we know that a charged particle in an electromagnetic field can be described by the substitution $p_\mu \to p_\mu - QeA_\mu$ where Q is the electric charge of the particle in units of e (for electrons, $Q=-1$). A quantum-mechanical analogue is to replace the operator $i\partial_\mu$ by $i\partial_\mu - QeA_\mu$. Substituting this in the DE, we get an interaction term

$$\mathcal{L}_{int} = -Qe\bar\psi\gamma^\mu\psi A_\mu\,. \tag{21}$$

This gives the entire quantum electrodynamics. Now note that physics remains invariant under the gauge transformation $A_\mu \to A_\mu + \partial_\mu\lambda(x)$, where $\lambda(x)$ is a scalar function of space-time. Taking the local phase transformation and the gauge transformation together, we see that a cancellation occurs and we get back the original Lagrangian (plus the interaction term) if and only if $\lambda = -\alpha/Qe$.

Thus, invariance under a local phase transformation necessitates a gauge field, and that is why such a transformation is generally called a *gauge transformation.* This is a local U(1) transformation; we will see more complicated examples later.

We often write $D_\mu = \partial_\mu + iQeA_\mu$. This is known as the *gauge covariant derivative*, or simply as the covariant derivative [3]. The structure of the covariant derivative is more complicated for a non-Abelian gauge group; electrodynamics is an abelian U(1).

Putting $Q=-1$, we can write the Lagrangian for an electron in an electromagnetic field as

$$\mathcal{L} = \bar\psi\left(i\gamma^\mu\partial_\mu + e\gamma^\mu A_\mu - m\right)\psi\,. \tag{22}$$

Q. In quantum field theory, the electromagnetic field A_μ is raised to the status of an operator. Calculate $[D_\mu, D_\nu]\psi$ and hence show that

$$[D_\mu, D_\nu] = iQeF_{\mu\nu}\,, \tag{23}$$

where $F_{\mu\nu}$ is the electromagnetic field tensor.

---

[3] There is another covariant derivative in General Relativity, and mathematically these two are related. Since we will never talk about that, there is no chance of confusion.

1.4 Free-particle solutions

We can Fourier expand the Dirac spinor

$$\psi_\alpha(t,\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}}\sqrt{\frac{1}{2E_p}}\sum_{s=\pm}\left[\mathbf{b}(\mathbf{p},s)u_\alpha(\mathbf{p},s)e^{-ipx} + \mathbf{d}^\dagger(\mathbf{p},s)v_\alpha(\mathbf{p},s)e^{ipx}\right], \tag{24}$$

where $\mathbf{b}$ and $\mathbf{d}$ are the annihilation operators for the particle and the antiparticle respectively; daggers denote corresponding creation operators.

The momentum space spinors (remember, they are *not* operators!) satisfy the Dirac equation

$$(\not{p} - m)u(\mathbf{p},s) = (\not{p} + m)v(\mathbf{p},s) = 0\,. \tag{25}$$

The solutions are

$$u(\mathbf{p},+) = N\begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m}\end{pmatrix},\, u(\mathbf{p},-) = N\begin{pmatrix} 0 \\ 1 \\ \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m}\end{pmatrix},\, v(\mathbf{p},+) = N\begin{pmatrix} \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \\ 1 \\ 0\end{pmatrix},\, v(\mathbf{p},-) = N\begin{pmatrix} \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \\ 0 \\ 1\end{pmatrix}, \tag{26}$$

with $N = \sqrt{E+m}$, and the spin quantization along the z-axis.

Q. Check that ψs are orthogonal: $\psi_r^\dagger\psi_s = 0$ if $r \neq s$.
Q. With a normalisation of $2E$ particles per unit volume, show that $N = \sqrt{E+m}$.

Q. Explicitly check that one can write

$$u_r(\mathbf{p})^\dagger u_s(\mathbf{p}) = 2E\delta_{rs}\,, \quad v_r(\mathbf{p})^\dagger v_s(\mathbf{p}) = 2E\delta_{rs}\,, \quad u_r(\mathbf{p})^\dagger v_s(-\mathbf{p}) = 0\,, \tag{27}$$

where $r, s = 1, 2$.

Q. Derive the completeness relations:

$$\sum_{s=1}^{2} u_s\bar{u}_s = \not{p} + m\,, \quad \sum_{s=1}^{2} v_s\bar{v}_s = -\not{p} + m\,. \tag{28}$$

While we have used the Dirac-Pauli representation, the completeness relations are representation-independent.

2 Discrete Symmetries

2.1 Parity

Parity or space inversion is a discrete symmetry that changes, in the passive viewpoint, a right-handed coordinate system into a left-handed one. In the active viewpoint, the ket $|\alpha\rangle$ is transformed under the parity operator P as $|\alpha\rangle \to P|\alpha\rangle$, so that the expectation value of $\mathbf{r}$ with respect to these two states are opposite in sign:

$$\langle\alpha|P^\dagger\mathbf{r}P|\alpha\rangle = -\langle\alpha|\mathbf{r}|\alpha\rangle, \tag{29}$$

which can be accomplished if $P^\dagger\mathbf{r}P = -\mathbf{r}$ or $\{\mathbf{r}, P\} = 0$.

Consider a position eigenket $|\mathbf{r}'\rangle$. P reverses the position coordinates, upto a phase factor $e^{i\delta}$ (δ real), so

$$P|\mathbf{r}'\rangle = e^{i\delta}|-\mathbf{r}'\rangle. \tag{30}$$

To prove this assertion, note that

$$\mathbf{r}P|\mathbf{r}'\rangle = -P\mathbf{r}|\mathbf{r}'\rangle = (-\mathbf{r}')P|\mathbf{r}'\rangle. \tag{31}$$

So $P|\mathbf{r}'\rangle$ is an eigenket of the position operator $\mathbf{r}$ with eigenvalue $-\mathbf{r}'$, so it must be the position eigenket $|-\mathbf{r}'\rangle$ upto a phase factor. By convention we take $e^{i\delta}=1$; so $P^2|\mathbf{r}'\rangle = |\mathbf{r}'\rangle$, or P^2 always has eigenvalue $+1$. Thus, $P^{-1} = P^\dagger = P$, with possible eigenvalues $+1$ and -1 (remember that all kets are eigenkets of P^2, but need not be of P; an example is the momentum eigenket $|\mathbf{p}\rangle$, whose wavefunction is just the plane wave: the reason is that P flips momentum). Just like $\mathbf{r}$, the momentum operator $\mathbf{p}$ also anticommutes with P. Physically, this means that translation (whose generator is momentum) followed by P is P followed by translation in the opposite direction.

Thus, $\mathbf{r}$ and $\mathbf{p}$ are odd under parity. Angular momentum, defined as $\mathbf{L}=\mathbf{r}\times\mathbf{p}$, is even: $[P,\vec{L}]=0$. In quantum mechanics, we postulate that $[P,\mathcal{U}_R]=0$, so $[P,\mathbf{J}]=0$ since $\mathbf{J}$ is the generator of rotation. But $\mathbf{J}=\mathbf{L}+\mathbf{S}$, so spin is also even under parity. From the discussion of the previous section, we saw that $\mathbf{r}$, $\mathbf{p}$, and $\mathbf{J}$ are all vectors, or spherical tensors of rank 1. But the first two are polar vectors while the last one is an axial vector.

The operators $\mathbf{L}.\mathbf{S}$ or $\mathbf{r}.\mathbf{p}$ are ordinary scalars. But the operator $\mathbf{S}.\mathbf{r}$ picks up a minus sign under parity:

$$P^{-1}\mathbf{S}.\mathbf{r}P = -\mathbf{S}.\mathbf{r}. \tag{32}$$

This transformation property defines a pseudoscalar operator.

2.2 Time reversal

Time reversal, or reversal of motion to be precise, is nothing but running a film backward. If the Hamiltonian is symmetric under time reversal, you cannot distinguish between forward and backward runnings of the film. Mathematically speaking, if $m\ddot{\mathbf{r}} = -\nabla V(\mathbf{r})$ has a solution $\mathbf{r}(t)$, then $\mathbf{r}(-t)$ is also a solution. Note that the force should not be a dissipative one, like friction!

For classical electrodynamics, time reversal means $\mathbf{E}\to\mathbf{E}$, $\mathbf{B}\to-\mathbf{B}$, $\rho\to\rho$, $\mathbf{j}\to-\mathbf{j}$, and $\mathbf{v}\to-\mathbf{v}$. This keeps Maxwell's equations as well as Lorentz force law invariant. This is also intuitively clear: since $t\to-t$, both current and velocity flip sign, and magnetic field is caused by tiny currents, so that also flips. (How should ϕ and $\mathbf{A}$ change?)

Consider the Schrödinger equation:

$$i\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi, \tag{33}$$

and let $\psi(\mathbf{r},t)$ be a solution. Obviously, $\psi(\mathbf{r},-t)$ is not a solution, since we have a first derivative in t. But take the complex conjugate; then $\psi^*(\mathbf{r},-t)$ is a solution. This is also intuitively obvious: both $\psi(\mathbf{r},t)$ and $\psi^*(\mathbf{r},-t)$ go like $\exp(-iEt/\hbar)$. Thus, time reversal has something to do with complex conjugation. The time-reversed wavefunction corresponding to $\psi(\mathbf{r},t)=\langle\mathbf{r}|\alpha\rangle$ is $\langle\mathbf{r}|\alpha\rangle^* = \langle\alpha|\mathbf{r}\rangle$.

For any symmetry operation, we expect the inner product to be preserved: if $|\alpha\rangle\to|\tilde{\alpha}\rangle$, $|\beta\rangle\to|\tilde{\beta}\rangle$, then $\langle\tilde{\alpha}|\tilde{\beta}\rangle = \langle\alpha|\beta\rangle$. This follows from the property of unitary operators: $\langle\tilde{\alpha}|\tilde{\beta}\rangle =$

$\langle\alpha|U^\dagger U|\beta\rangle = \langle\alpha|\beta\rangle$. However, for time reversal we cannot take the symmetry operation to be unitary (we'll see later why), and we must impose a weaker condition: $|\langle\tilde{\alpha}|\tilde{\beta}\rangle| = |\langle\alpha|\beta\rangle|$. This admits another possibility, namely, $\langle\tilde{\alpha}|\tilde{\beta}\rangle = \langle\alpha|\beta\rangle^* = \langle\beta|\alpha\rangle$ (*i.e.*, ket and bra are interchanged).

Consider an operator θ. If this is antiunitary, then for the transformations $|\alpha\rangle \to |\tilde{\alpha}\rangle = \theta|\alpha\rangle$, $|\beta\rangle \to |\tilde{\beta}\rangle = \theta|\beta\rangle$, we must have

$$\langle\tilde{\alpha}|\tilde{\beta}\rangle = \langle\alpha|\beta\rangle^*, \quad \theta\left(c_1|\alpha\rangle + c_2|\beta\rangle\right) = c_1^*\theta|\alpha\rangle + c_2^*\theta|\beta\rangle. \tag{34}$$

The last equation alone defines an antilinear transformation.

Any antiunitary operator can be written as $\theta = UK$ where U is a unitary operator and K is the complex conjugation operator defined as $Kc|\alpha\rangle = c^*K|\alpha\rangle$.

The proof is easy. Putting $\theta = UK$ in the second equation of (34), we get

$$UK\left(c_1|\alpha\rangle + c_2|\beta\rangle\right) = c_1^*UK|\alpha\rangle + c_2^*UK|\beta\rangle = c_1^*\theta|\alpha\rangle + c_2^*\theta|\beta\rangle. \tag{35}$$

Before proving the first relation of (34), let us see what $\theta|\alpha\rangle$ is. $|\alpha\rangle$ can be expanded in terms of the basis kets $|i\rangle$: $|\alpha\rangle = \sum_i |i\rangle\langle i|\alpha\rangle$. But $|i\rangle$ is a column vector with one 1 and rest 0, so $K|i\rangle = |i\rangle$. Thus

$$|\alpha\rangle \to |\tilde{\alpha}\rangle = \sum_i \langle i|\alpha\rangle^* UK|i\rangle = \sum_i \langle\alpha|i\rangle U|i\rangle. \tag{36}$$

Similarly, $|\tilde{\beta}\rangle = \sum_j \langle\beta|j\rangle U|j\rangle$, so $\langle\tilde{\beta}| = \sum_j \langle j|\beta\rangle\langle j|U^\dagger$. Thus,

$$\langle\tilde{\beta}|\tilde{\alpha}\rangle = \sum_{i,j}\langle j|\beta\rangle\langle j|U^\dagger U|i\rangle\langle\alpha|i\rangle = \sum_i \langle\alpha|i\rangle\langle i|\beta\rangle = \langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*. \tag{37}$$

2.3 Charge conjugation

The charge conjugation operator C does not affect the space-time coordinates but flips the charge of the particle. Consider the non-relativistic Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m}, \tag{38}$$

which, under the minimal substitution $p^\mu \to p^\mu - eA^\mu$, goes to

$$H = \frac{1}{2m}\left[\mathbf{p}^2 - e(\mathbf{p}.\mathbf{A} + \mathbf{A}.\mathbf{p}) + e^2\mathbf{A}^2\right] + e\phi\,. \tag{39}$$

If this is to remain invariant under C which takes $e \to -e$, we must have $A^\mu \to -A^\mu$; thus, electromagnetic field is odd under C. The Hamiltonian is invariant, *i.e.*, $CHC^{-1} = H$ or $[C, H] = 0$. Also, C^2 takes a fieldback to itself, so it must have an eigenvalue $\exp(2i\delta)\mathbf{1}$, just like parity, and by convention we take $\delta = 0$.

2.4 Symmetries in field theory

The symmetry operators should leave all physical observables invariant. In particular, we would expect that under any symmetry operation $\mathcal{S}$,

- The ground state or vacuum remains invariant, $\mathcal{S}|0\rangle = |0\rangle$.

- The action remains invariant, $\int \mathcal{L} d^4x \to \int \mathcal{L} d^4x$. In other words, $[\mathcal{S}, H] = 0$.
- The quantization conditions remain invariant. Of course, if this does not happen, we can't do anything with the transformed theory. In ordinary quantum mechanics, this means having a transformation under which $[x, p] = i$ is no longer valid.

If the first condition is not satisfied, we call that a *spontaneous* breaking of symmetry; the action is still invariant but the vacuum is no longer so. If the second condition is violated, it is an *explicit* breaking of symmetry. Both discrete and spontaneous symmetries can be broken either spontaneously or explicitly.

Q. Suppose under some symmetry operation $\mathcal{S}$ the magnitude squared of the scalar product of two vectors remain invariant. Show that $\mathcal{S}$ must be either unitary or antiunitary. Show that for a unitary transformation, the commutation relation $[x, p] = i$ remains valid.

2.5 Spin-1 fields

From our experience with classical fields, we can predict

$$PA_\mu(t, \mathbf{x})P^{-1} = A^\mu(t, -\mathbf{x}) \,, \quad CA_\mu(t, \mathbf{x})C^{-1} = -A_\mu(t, \mathbf{x}) \,, \quad TA_\mu(t, \mathbf{x})T^{-1} = A^\mu(-t, \mathbf{x}) \,. \tag{40}$$

Let us try to see intuitively how they come. The transformation of the arguments is obvious, that follows from the definition of these operators. Take A^μ to be the electromagnetic four-potential (in QFT, this is then the photon field). Under parity, ϕ does not change but $\mathbf{A}$, being an ordinary vector, picks up a minus sign, so $A_\mu \to A^\mu$ in the Bjorken-Drell metric. Under T, again, $\mathbf{A} \to -\mathbf{A}$, as the vector potential comes because of the current, and the direction of the current flips under T. For C, as we have just seen, the overall minus sign is necessary to keep electromagnetism invariant under parity.

Q. If the electromagnetic Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - eJ_\mu A^\mu \tag{41}$$

is invariant under C, P and T separately, show that J_μ must have identical transformation laws as A_μ.

Q. How does ∂_μ change under P and T? Hence show that

$$P\mathcal{L}(t, \mathbf{x})P^{-1} = \mathcal{L}(t, -\mathbf{x}) \,, \quad C\mathcal{L}(t, \mathbf{x})C^{-1} = \mathcal{L}(t, \mathbf{x}) \,, \quad T\mathcal{L}(t, \mathbf{x})T^{-1} = \mathcal{L}(-t, \mathbf{x}) \,. \tag{42}$$

Hence argue that the action will remain invariant, even though the arguments of $\mathcal{L}$ might change.

Q($*$) Expand the photon field in terms of the annihilation operator $\mathbf{a}$ and creation operator $\mathbf{a}^\dagger$. Remember that the components of the polarization vector ϵ_μ are constant numbers and are not affected by P or C. However, T changes an incoming photon to an outgoing photon, so

$$T\epsilon^\mu_\pm T^{-1} = (\epsilon^\mu_\pm)^* = \epsilon^\mu_\mp \,, \tag{43}$$

where ϵ_+ and ϵ_- are two polarization (spin) degrees of freedom. Following the transformation of A^μ under C, P and T, and suitably changing the arguments, show that

$$P\mathbf{a}(\mathbf{p}, \pm)P^{-1} = -\mathbf{a}(-\mathbf{p}, \pm) \,, \quad C\mathbf{a}(\mathbf{p}, \pm)C^{-1} = -\mathbf{a}(\mathbf{p}, \pm) \,, \quad T\mathbf{a}(\mathbf{p}, \pm)T^{-1} = -\mathbf{a}(-\mathbf{p}, \mp) \,. \tag{44}$$

Thus, the *intrinsic parity* and *C-parity* of a one-photon state is negative:

$$P|\gamma; \mathbf{p}\rangle = -|\gamma; -\mathbf{p}\rangle \,, \quad C|\gamma; \mathbf{p}\rangle = -|\gamma; \mathbf{p}\rangle \,. \tag{45}$$

2.6 Spin-$\frac{1}{2}$ fields: parity

It is not yet clear why γ^μ, a set of four matrices, is treated as a contravariant Lorentz vector. However, consider a Lorentz transformation $x'^\mu = \Lambda^\mu_\nu x^\nu$, under which the Dirac spinor $\psi(x)$ transforms linearly: $\psi(x) \to \psi'(x') = S(\Lambda)\psi(x)$. Under the same transformation, the derivative changes as

$$\partial_\mu \to \partial'_\nu = \Lambda^\mu_\nu \partial_\mu \,, \tag{46}$$

so the Dirac equation in the boosted frame looks as

$$\left(i\gamma^\nu \partial'_\nu - m\right)\psi'(x') = \left(i\gamma^\nu \Lambda^\mu_\nu \partial_\mu - m\right) S(\Lambda)\psi(x)\,. \tag{47}$$

If we multiply (47) by the matrix S^{-1} from right, it is easy to see that the Dirac equation remains invariant if and only if

$$S^{-1}\gamma^\nu \Lambda^\mu_\nu S = \gamma_\mu \Longrightarrow S\gamma^\mu S^{-1} = \Lambda^\mu_\nu \gamma^\nu\,. \tag{48}$$

This is precisely why the set of matrices γ^μ is considered to be a contravariant Lorentz vector, though none of them are functions of x. In other words, it is the combination $\bar\psi\gamma^\mu\psi$ that behaves as a contravariant 4-vector.

While the parity transformation $t \to t, \mathbf{x} \to -\mathbf{x}$ is an improper Lorentz transformation, we can still write $\Lambda^\mu_\nu = \mathrm{diag}(1, -1, -1, -1)$. Thus, the matrix S should satisfy

$$S\gamma^0 = \gamma^0 S\,, \quad S\gamma^i = -\gamma^i S\,, \tag{49}$$

which is satisfied by $S = \gamma^0$ (or any multiple of γ^0). Thus, $\psi_P(t, \mathbf{x}) \equiv P\psi(t, \mathbf{x})P^{-1} = \gamma^0\psi(t, -\mathbf{x})$ is the parity-transformed spinor. In the Dirac-Pauli representation, $\gamma^0 = \mathrm{diag}(1, 1, -1, -1)$; so

$$(\psi_P)_i = \psi_i \quad (i = 1, 2) \qquad (\psi_P)_i = -\psi_i \quad (i = 3, 4)\,. \tag{50}$$

Thus, the spinors are eigenspinors of parity, but *particle and antiparticle spinors have opposite parity eigenvalues.* This eigenvalue is known as the *intrinsic parity*; while the choice of +1 for particle is arbitrary, the eigenvalues for particle and antiparticle spinors must be opposite.

One can easily check that

$$u(-\mathbf{p}, s) = \gamma^0 u(\mathbf{p}, s)\,, \quad v(-\mathbf{p}, s) = -\gamma^0 v(\mathbf{p}, s)\,. \tag{51}$$

Now, one can write

$$\begin{aligned}
&P\psi(t, \mathbf{x})P^{-1} \\
= &\int \frac{d^3\mathbf{p}}{\sqrt{(2\pi)^3 2E_p}} \sum_{s=\pm} \left[P\mathbf{b}(\mathbf{p}, s)P^{-1}u(\mathbf{p}, s)e^{-ipx} + P\mathbf{d}^\dagger(\mathbf{p}, s)P^{-1}v(\mathbf{p}, s)e^{ipx}\right] \\
= &\int \frac{d^3\mathbf{p}}{\sqrt{(2\pi)^3 2E_p}} \sum_{s=\pm} \left[P\mathbf{b}(-\mathbf{p}, s)P^{-1}u(-\mathbf{p}, s)e^{-i(E_pt+\mathbf{p}.\mathbf{x})} + P\mathbf{d}^\dagger(-\mathbf{p}, s)P^{-1}v(-\mathbf{p}, s)e^{i(E_pt+\mathbf{p}.\mathbf{x})}\right] \\
= &\int \frac{d^3\mathbf{p}}{\sqrt{(2\pi)^3 2E_p}} \sum_{s=\pm} \left[P\mathbf{b}(-\mathbf{p}, s)P^{-1}\gamma^0 u(\mathbf{p}, s)e^{-i(E_pt+\mathbf{p}.\mathbf{x})} - P\mathbf{d}^\dagger(-\mathbf{p}, s)P^{-1}\gamma^0 v(\mathbf{p}, s)e^{i(E_pt+\mathbf{p}.\mathbf{x})}\right]
\end{aligned} \tag{52}$$

where in the second step we have changed the variable $\mathbf{p} \to -\mathbf{p}$ (and $d^3\mathbf{p} \to -d^3\mathbf{p}$, with integration limits flipped).

Q. In the Weyl representation, $\gamma^0 = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}$. Show that under the parity transformation, ψ_L goes to ψ_R and vice versa.

Q. Identifying eq. (52) with $\gamma^0\psi(t, -\mathbf{x})$, show that

$$P\mathbf{b}(\mathbf{p}, s)P^{-1} = \mathbf{b}(-\mathbf{p}, s)\,, \quad P\mathbf{d}(\mathbf{p}, s)P^{-1} = -\mathbf{d}(-\mathbf{p}, s)\,, \tag{53}$$

which again shows that the intrinsic parities of a fermion and an antifermion are opposite.

Q. A meson (or a positronium) is a fermion-antifermion bound state. Parity of this state is defined as the orbital parity times the intrinsic parities of the two components. Show that the parity eigenvalue is $-(-1)^L$, where L is the orbital angular momentum. Why doesn't the spin play a part here?

Q. Show that the anticommutation relations remain the same for ψ_P, *i.e.*, show that

$$\{\psi_{\alpha P}(t, \mathbf{x}), \psi^\dagger_{\beta P}(t, \mathbf{y})\} = \delta^{(3)}(\mathbf{x} - \mathbf{y})\delta_{\alpha\beta}\,. \tag{54}$$

2.7 Spin-$\frac{1}{2}$ fields: charge conjugation

The charge conjugation operator C transforms an electron spinor to a positron spinor and vice versa. The electron wavefunction satisfies

$$(i\gamma^\mu\partial_\mu + e\gamma^\mu A_\mu - m)\psi_{e^-} = 0\,, \tag{55}$$

which follows from the minimal substitution $\partial_\mu \to \partial_\mu + iQeA_\mu$, with $Q = -1$ for electron, and e is the charge of the proton. The positron wavefunction must satisfy

$$(i\gamma^\mu\partial_\mu - e\gamma^\mu A_\mu - m)\psi_{e^+} = 0\,. \tag{56}$$

The question is: how are these two wavefunctions related? Taking the complex conjugate of (55) won't do, because the signs are not all correct:

$$(i\gamma^{\mu *}\partial_\mu - e\gamma^{\mu *}A_\mu + m)\psi^*_{e^-} = 0\,. \tag{57}$$

However, if we multiply (57) from left by a matrix $C\gamma^0$ and demand $C\gamma^0\gamma^{\mu *} = -\gamma^\mu C\gamma^0$, we are through:

$$(i\gamma^\mu\partial_\mu - e\gamma^\mu A_\mu - m)C\gamma^0\psi^*_{e^-} = 0\,, \tag{58}$$

and we can identify

$$\psi_{e^+} = C\gamma^0\psi^*_{e^-}\,. \tag{59}$$

We often simply write this as

$$\psi_C = C\gamma^0\psi^* = C\bar{\psi}^T\,. \tag{60}$$

The complex conjugation changes the $\exp(-ipx)$ spinor to the $\exp(ipx)$ one, which we identify as the antiparticle spinor. In the Dirac-Pauli representation, it is easy to show that the choice $C = i\gamma^2\gamma^0$ satisfies $C\gamma^0\gamma^{\mu *} = -\gamma^\mu C\gamma^0$.

Note that $(\psi_C)_L = (\psi_R)_C$; this is easy to see, as

$$(\psi_C)_L = P_L\psi_C = P_L C\gamma^0\psi^* = C\gamma^0 P_R\psi^* = (\psi_R)_C\,. \tag{61}$$

Q. In the Dirac-Pauli representation, show that

$$C^{-1} = C^\dagger = C^T = -C\,, \quad C^{-1}\gamma^\mu C = -\gamma^{\mu T}\,, \quad \overline{\psi_C} = -\psi^T C^{-1}\,. \tag{62}$$

2.8 Spin-$\frac{1}{2}$ fields: time reversal

The time reversal operator $T = UK$ should be such that

$$\psi^T(t,\mathbf{x}) = T\psi(t,\mathbf{x})T^{-1} = U\psi(t,-\mathbf{x})\,. \tag{63}$$

The electromagnetic current transforms as

$$\begin{aligned} TJ^\mu(t,\mathbf{x})T^{-1} &= T\bar\psi(t,\mathbf{x})\gamma^\mu\psi(t,\mathbf{x})T^{-1} \\ &= T\bar\psi(t,\mathbf{x})T^{-1}\gamma^{\mu *}T\psi(t,\mathbf{x})T^{-1} = \bar\psi(-t,\mathbf{x})U^\dagger\gamma^{\mu *}U\psi(-t,\mathbf{x})\,. \end{aligned} \tag{64}$$

So T should be such that

$$U^\dagger\gamma^{\mu *}U = \gamma_\mu\,. \tag{65}$$

In the Dirac-Pauli representation, it is easy to check that $U = \gamma^1\gamma^3$ has this property.

Q. Using the explicit form of the momentum-space spinors, show that

$$\gamma^1\gamma^3 u(\mathbf{p},s)^* = -su(-\mathbf{p},-s)\,,\quad \gamma^1\gamma^3 v(\mathbf{p},s)^* = -su(-\mathbf{p},-s)\,. \tag{66}$$

Q. Show that under T, a $V-A$ current retains its form.

2.9 Bilinear covariants

The Dirac Lagrangian

$$\mathcal{L} = i\bar\psi(\gamma^\mu\partial_\mu - m)\psi \tag{67}$$

contains two quantities: $\bar\psi\gamma^\mu\psi$ and $\bar\psi\psi$. These quantities are called *bilinear covariants* or simply bilinears; the name stems from the fact that ψ and $\bar\psi$ have to be treated as independent variables in obtaining the Dirac equation from the Lagrangian. A quantity is called *covariant* when it transforms according to some representation of the Lorentz group. Depending on the number of γ-matrices sandwiched between $\bar\psi$ and ψ, we can form 16 independent bilinear covariants. The third column shows the number of these covariants. Note that $\sigma^{\mu\nu}$ is an antisymmetric tensor and hence there are only 6 of them. Note that γ^μ is *not* a Lorentz vector; they are constant matrices and do not transform under coordinate transformations. We put a Lorentz index because $\bar\psi\gamma^\mu\psi$ transforms like a contravariant 4-vector.

Scalar (S)	$\bar\psi\psi$	1
Vector (V)	$\bar\psi\gamma^\mu\psi$	4
Tensor (T)	$\bar\psi\sigma^{\mu\nu}\psi$	6
Axialvector (A)	$\bar\psi\gamma^\mu\gamma_5\psi$	4
Pseudoscalar (P)	$\bar\psi\gamma_5\psi$	1

Q. Show that $\bar\psi\sigma^{\mu\nu}\gamma_5\psi$ is not an independent bilinear covariant. More generally, show that the sixteen 4×4 matrices $\Gamma = \{\mathbf{1}, \gamma^\mu, \sigma^{\mu\nu}, \gamma^\mu\gamma_5, \gamma_5\}$ completely span the vector space formed by all 4×4 matrices. For this, what you have to show is: the equation $\sum_i a_i\Gamma_i = 0$ ($i = 1$ to 16) for complex coefficients a_i can be satisfied if and only if all a_is are zero. [Hint: Multiply by each of the Γ_is in turn and take the trace, use the trace theorems.]

It is important to check the transformation of these bilinears under the discrete symmetries discussed above. Let us start with parity. Note that $\bar\psi_P = \bar\psi\gamma^0$.

$$\begin{aligned}\bar\psi_P\psi_P &= \bar\psi\gamma^0\gamma^0\psi = \bar\psi\psi\,,\\ \bar\psi_P\gamma_5\psi_P &= \bar\psi\gamma^0\gamma_5\gamma^0\psi = -\bar\psi\gamma_5\psi\,,\end{aligned}\tag{68}$$

which shows why the second one is called a pseudoscalar. For vector bilinears, we have to be a bit more careful and remember that while $\gamma^0\gamma^0\gamma^0 = \gamma^0$, $\gamma^0\gamma^i\gamma^0 = -\gamma^i$, so that

$$\begin{aligned}\bar\psi_P\gamma^\mu\psi_P &= \bar\psi\gamma^0\gamma^\mu\gamma^0\psi = \bar\psi\gamma_\mu\psi\,,\\ \bar\psi_P\gamma^\mu\gamma_5\psi_P &= \bar\psi\gamma^0\gamma^\mu\gamma_5\gamma^0\psi = -\bar\psi\gamma_\mu\gamma_5\psi\,.\end{aligned}\tag{69}$$

The Lagrangian has to be a Lorentz scalar and so the vector (or axialvector) current has to be contracted with another Lorentz vector A_μ. Under parity, its time component does not change sign, but the spatial components do, so it becomes A^μ, and the contraction is again properly defined.

Q. Find the parity transformation properties of σ^{0i} and σ^{ij}.

Q. Show that the $V-A$ current $\bar\psi\gamma^\mu(1-\gamma_5)\psi$ becomes a $V+A$ current under parity.

For charge conjugation, we use the results in (62). For example,

$$\bar\psi_C\psi_C = -\psi^T C^{-1}C\bar\psi^T = -\psi^T\bar\psi^T = \bar\psi\psi\,.\tag{70}$$

There are two important steps that have gone in (70). First, $\psi^T\bar\psi^T$ is a 1×1 object, so the transpose is equal to the original function. Second, and the more important, is an extra minus sign in the last step that makes the overall sign positive. The origin of this can be understood only when we go into the quantisation of Dirac fields. To pre-empt, let us just say that the wavefunctions are raised to the status of operators, and the fermionic operators ψ and $\bar\psi$ anticommute. So changing their positions brings in an extra minus sign.

Similarly,

$$\bar\psi_C\gamma^\mu\psi_C = -\psi^T C^{-1}\gamma^\mu C\bar\psi^T = -\psi^T(-\gamma^\mu)^T\bar\psi^T = -\bar\psi\gamma^\mu\psi\,.\tag{71}$$

This shows that the fermionic current $\bar\psi\gamma^\mu\psi$ picks up a minus sign under charge conjugation; so the electron and positron currents have opposite sign, which we may also interpret as opposite couplings to the electromagnetic field. Again, this is nothing but the statement that the electric charges of electron and positron are opposite.

Q. Show that $\bar\psi_C\gamma_5\psi_C = \bar\psi\gamma_5\psi$ and $\bar\psi_C\gamma^\mu\gamma_5\psi_C = \bar\psi\gamma^\mu\gamma_5\psi$.

Q. Hence show that the $V-A$ current changes to $V+A$ under charge conjugation.

Q. Convince yourself that under the combined operation CP (or PC, the order is irrelevant) the $V-A$ current retains its form.

Q. Show that $Tr[\gamma^\alpha\gamma^\beta\cdots\gamma^\mu\gamma^\nu] = Tr[\gamma^\nu\gamma^\mu\cdots\gamma^\beta\gamma^\alpha]$, *i.e.*, the order of the γ-matrices inside a trace can be reversed. [Hint: Consider the matrix $D = -iC = \gamma^2\gamma^0$. Obviously $D^2 = \mathbf{1}$, so that $D = D^{-1}$, and $D^{-1}\gamma^\mu D = (\gamma^\mu)^T$. Now write the trace as $Tr[D^2\gamma^\mu D^2\gamma^\nu\cdots D^2\gamma^\alpha D^2\gamma^\beta] = Tr[D^{-1}\gamma^\mu DD^{-1}\gamma^\nu D\cdots D^{-1}\gamma^\alpha DD^{-1}\gamma^\beta D]$. Show that this is just $(-1)^n$ times the trace of the order of the γ-matrices reversed, where n is the number of the γ-matrices inside the trace. Again, n must be even for the trace to be nonzero.]

A warning: T has been used in this section in at least three different contexts: time reversal operator, transpose of a matrix, and tensor bilinear. It is expected that the reader can guess which T is what.

3 SU(3) and the Quark Model

3.1 SU(3)

$SU(3)$ has eight generators, and two of them can be simultaneously diagonalised at most; this is a rank-2 group. The generator matrices are conventionally written as $X_a = \frac{1}{2}\lambda_a$, a running from 1 to 8. The Gell-Mann matrices λ_a are as follows:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{72}$$

Note that all matrices are traceless and hermitian, and just like the fundamental representation of $SU(2)$, the generators are diagonalised according to $Tr(X_i X_j) = \frac{1}{2}\delta_{ij}$.

The fundamental representation **3** is complex, *i.e.*, inequivalent to $\bar{\mathbf{3}}$. The states are labeled by the weights, the eigenvalues corresponding to X_3 and X_8. Note that λ_1, λ_2, and λ_3 constitute an $SU(2)$ in the 1-2 block.

The 3×3 λ matrices satisfy the following relations:

$$[\lambda_a, \lambda_b] = 2if_{abc}\lambda_c, \quad \{\lambda_a, \lambda_b\} = \frac{4}{3}\delta_{ab}\mathbf{1} + 2d_{abc}\lambda_c, \quad Tr(\lambda_a\lambda_b) = 2\delta_{ab}. \tag{73}$$

The completeness relation is

$$(\lambda_a)_{ij}(\lambda_b)_{kl} = -\frac{2}{3}\delta_{ij}\delta_{kl} + 2\delta_{il}\delta_{jk} \tag{74}$$

where $i, j, k, l = 1, 2, 3$. The quadratic and cubic Casimir operators are defined as

$$C_2(R)\mathbf{1} \equiv \sum_{a=1}^{8}(X_a)^2(R), \quad C_3(R)\mathbf{1} \equiv \sum_{a,b,c=1}^{8} d_{abc}X_a(R)X_b(R)X_c(R), \tag{75}$$

where R is the dimension of the representation (3 for fundamental) and $X_a = \lambda_a/2$. Any representation of dimension d can be denoted by a set of two non-negative integers (p, q), with

$$d = \frac{1}{2}(p+1)(q+1)(p+q+2)\,. \tag{76}$$

For $(p, q) = (0, 0)$, we have a singlet, **1**; $(1, 0)$ is the fundamental **3**. If $p = q$, the representation is real, otherwise it is complex. If $q > p$, we put a bar on the representation to denote that it is the complex conjugated representation of (p, q); thus, $(0, 1)$ is $\bar{\mathbf{3}}$.

It can be shown that

$$\begin{aligned} C_2(p,q) &= \frac{1}{3}\left[p^2 + pq + q^2 + 3p + 3q\right], \\ C_3(p,q) &= \frac{1}{18}(p-q)(2p+q+3)(2q+p+3). \end{aligned} \tag{77}$$

This gives, *e.g.*, $C_2(3) = 4/3$, $C_2(8) = 3$.

Now, some important relations. From eq. (74), we have

$$(\lambda_a)_{ij}(\lambda_a)_{jl} = \left(-\frac{2}{3} + 2 \times 3\right)\delta_{il} = 4C_2(3)\delta_{il}. \tag{78}$$

From the definition of the adjoint representation $(X_a)_{bc} = -if_{abc}$,

$$f_{acd}f_{bcd} = C_2(8)\delta_{ab} = 3\delta_{ab}. \tag{79}$$

Q. Show that $f_{abc}\lambda_b\lambda_c = iC_2(8)\lambda_a$. (*Hint*: write $\lambda_b\lambda_c = \frac{1}{2}[\lambda_b, \lambda_c]$ since f_{abc} is completely antisymmetric, and then use eq. (79).)

Q. Justify the steps:

$$\begin{aligned} \lambda_b\lambda_a\lambda_b &= \frac{1}{2}\left(\lambda_b[\lambda_a, \lambda_b] - [\lambda_a, \lambda_b]\lambda_b + \lambda_b\lambda_b\lambda_a + \lambda_a\lambda_b\lambda_b\right) \\ &= 4C_2(3)\lambda_a + if_{abc}[\lambda_b, \lambda_c] = 4\left\{C_2(3) - \frac{1}{2}C_2(8)\right\}\lambda_a. \end{aligned} \tag{80}$$

3.2 The quark model

We have already mastered the basic mathematical tools necessary for an understanding of the celebrated idea of Gell-Mann and Nishijima about point-like constituents inside mesons and baryons, called quarks, and their antiparticles, called antiquarks. How do we know that there are indeed such constituents inside the hadrons? That followed from a celebrated experiment, one that we will discuss later, but in essence it is similar to the Rutherford experiment that showed point-like nucleus inside an atom: abnormal deviation of the projectiles, in this case electrons. There is another no less compelling argument. A neutral pointlike particle must have zero magnetic moment, as it cannot interact with the electromagnetic field. Neutrons, however, have a large negative magnetic moment. The magnetic moment of protons, too, is not what you expect for a particle of mass 938 MeV and unit positive charge. So they cannot be pointlike objects; there must be some internal structure. Are these internal structures themselves point-like, or is there a hierarchy going on? The answer is that we do not know yet. There may be a substructure for quarks; higher energy colliders will hopefully reveal that. For the time being, our knowledge ends at quarks as point-like, and hence structureless, objects.

A quark and an antiquark form a meson; three quarks form a baryon, and three antiquarks form an antibaryon. Why only these combinations? The answer comes from the underlying gauge theory of strong interactions, called Quantum Chromodynamics (QCD). The gauge group of QCD is SU(3), and it is unbroken; the symmetry is an exact one. So just like U(1) of electromagnetism, the gauge bosons — now they are 8 in number, equal to the number of generators of the gauge group — are massless; they are called gluons, and live in the 8-dimensional adjoint representation of SU(3). Each quark lives in a fundamental representation of SU(3), which is 3-dimensional, so each quark comes with three different quantum numbers. We call them *colour*, and let the three colours be denoted by R (red), G (green) and B (blue), after the three basic colours in optics. The SU(3) is exact, so we cannot distinguish between the colours; it is impossible from the experiments to say whether an up quark is u_R or u_G or u_B. The antiquarks carry anticolours, denoted by $\bar{\text{R}}$, $\bar{\text{G}}$, and $\bar{\text{B}}$. Only those particles which transform nontrivially under this SU(3) can take part in strong interaction; leptons are singlets and hence do not feel strong interaction.

We will discuss later in this chapter why the idea of colour was at all needed. But the theory predicts that every observable hadron must have a singlet colour wavefunction [4]. With $\mathbf{3}$ and $\bar{\mathbf{3}}$, there are only two ways to get the singlet:

$$\mathbf{3}\otimes\bar{\mathbf{3}} = \mathbf{8}\oplus\mathbf{1}\,, \quad \mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3} = \mathbf{10}\oplus\mathbf{8}\oplus\mathbf{8}\oplus\mathbf{1}\,. \tag{81}$$

So only these combinations (plus three $\bar{\mathbf{3}}$s for antibaryons) are allowed.

There is, however, another SU(3) that is more relevant for the construction of hadrons. The different types of quarks are called *flavours*, and that is why this SU(3) is called the SU(3) of flavour and written as $SU(3)_f$ (in contrast, the earlier SU(3) is called $SU(3)_c$ for colour). Before we go into that, let us focus on something more familiar.

3.3 Strong isospin

Suppose we have only two quarks, up (u) and down (d). The Lagrangian is

$$\mathcal{L} = i\left(\bar{u}\partial\!\!\!/ u + \bar{d}\partial\!\!\!/ d\right) - m_u\bar{u}u - m_d\bar{d}d\,. \tag{82}$$

The kinetic term remains invariant if we rotate the (u,d) basis to some other (u',d') basis given by

$$u' = u\cos\theta + d\sin\theta\,, \quad d' = -u\sin\theta + d\cos\theta\,. \tag{83}$$

The mass term breaks this symmetry. The symmetry is restored if we put $m_u = m_d = m$, since u' and d' will also have mass m. The introduction of electromagnetic interaction ultimately breaks this symmetry, because the electric charges of the quarks are different.

Experimentally, proton and neutron are almost mass degenerate. This motivated Heisenberg to propose the so-called *isospin symmetry* between p and n: this is an exact SU(2) as long as $m_n = m_p$ and very mildly broken by the mass difference ~ 1 MeV, which is thought to be electromagnetic in origin. The valence quark contents [5] of proton and neutron are uud and udd respectively, so the isospin symmetry can be translated to the u-d sector: $m_u = m_d$. The symmetry is almost exact, but not precisely. If it were a gauge symmetry, the gauge bosons would not have been exactly massless. We will call this *strong isospin* since there is another *weak isospin* in the theory of weak interaction, another SU(2) [6].

The fundamental representation of this SU(2) is $\begin{pmatrix} p \\ n \end{pmatrix}$, or $\begin{pmatrix} u \\ d \end{pmatrix}$. There is one diagonal generator — SU(2) is a rank-1 group — and so we need one quantum number to distinguish between these two states. That quantum number is taken to be the diagonal entries of $\frac{1}{2}\sigma_3$, *i.e.*, $\pm\frac{1}{2}$; $\frac{1}{2}$ for u and $-\frac{1}{2}$ for d. We say that u and d (or p and n) have isospin $I = \frac{1}{2}$ with projections $I_3 = \frac{1}{2}$ and $I_3 = -\frac{1}{2}$ respectively. All other quarks have isospin zero.

Two nucleons can be in an $I = 1$ or $I = 0$ combination. The symmetric $I = 1$ combinations are pp ($I_3 = 1$), $(pn + np)/\sqrt{2}$ ($I_3 = 0$), and nn ($I_3 = -1$). The antisymmetric $I = 0$ combination must have $I_3 = 0$ and hence orthogonal to $I = 1, I_3 = 0$: it is $(pn - np)/\sqrt{2}$. The nuclear force is such that the $I = 1$ combinations are unstable, but the $I = 0$, deuteron, is just stable.

---

[4] The protons and neutrons are bound inside a nucleus by the nuclear force, which is a remnant of the strong interaction, but that does not mean that protons and neutrons are not singlets under SU(3).

[5] The valence quarks are just the constituents of the hadrons. There are other quarks and antiquarks inside a hadron about which we will talk later.

[6] The name is a misnomer and a hopeless legacy of the original isotopic spin.

Antiquarks live in $\bar{\mathbf{2}}$; but SU(2) representations are all real and $\bar{\mathbf{2}}$ is equivalent to $\mathbf{2}$. However, $\begin{pmatrix} \bar{u} \\ \bar{d} \end{pmatrix}$ is not a correct assignment; the upper member must have $I_3 = \frac{1}{2}$, but I_3 of $\bar{u}$ is $-\frac{1}{2}$ (all such additive quantum numbers flip sign for antiparticles). Thus, $\begin{pmatrix} \bar{d} \\ \bar{u} \end{pmatrix}$ might be a correct assignment, but let us just check the transformation properties. For simplicity, consider a rotation about the isospin-y axis by an angle π. We will use

$$\exp\left[i\frac{\sigma_i}{2}\theta\right] = \cos\frac{\theta}{2} + i\sigma_i \sin\frac{\theta}{2}\,, \tag{84}$$

where σ_i is any of the three Pauli matrices. Thus, $\exp(i\pi\sigma_2/2) = i\sigma_2$, and

$$\begin{pmatrix} u' \\ d' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u \\ d \end{pmatrix} \Longrightarrow u' = -d\,,\quad d' = u\,. \tag{85}$$

There is nothing to worry about electric charge, as we have not yet introduced electromagnetism. Now we must have $\bar{u}' = -\bar{d}$ and $\bar{d}' = \bar{u}$. With the same transformation matrix, this is possible only if we define the antiparticle doublet as $\begin{pmatrix} -\bar{d} \\ \bar{u} \end{pmatrix}$ (there should be a relative minus sign, the position is not important). Now we are ready to define the combinations:

$$|\pi^+\rangle = |I=1, I_3=1\rangle = -u\bar{d}\,,\quad |\pi^0\rangle = |1,0\rangle = \frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})\,,\quad |\pi^-\rangle = |1,-1\rangle = d\bar{u}\,, \tag{86}$$

and an orthogonal $|0,0\rangle = (u\bar{u} + d\bar{d})/\sqrt{2}$. Due to the almost exact nature of strong isospin, the pions are almost mass degenerate: $m_{\pi^+} \approx m_{\pi^0}$. Of course, $m_{\pi^+} = m_{\pi^-}$ is a direct consequence of the CPT theorem.

Q. Prove (84).

3.4 The SU(3) of flavour

Now let us introduce the third quark, the strange (s). It is an unstable quark and decays via weak interaction only, and all mesons and baryons containing s are unstable too. Remember that only these three quarks were known in the 1960s, and Gell-Mann's original quark model was based on these three flavours. More flavours were added later as they were discovered.

Strange mesons and baryons are much heavier than their nonstrange counterparts. For example, one of the lowest-lying strange mesons, $K^-(= s\bar{u})$ is more than three times heavier than π^-. The immediate conclusion is that the s quark must be heavier than either u or d, and SU(3)$_f$ is a broken symmetry, broken by the mass difference $m_s - m_{u,d}$. Such a breaking is known as explicit breaking of a symmetry.

SU(3) is a rank-2 group with two diagonal generators $\frac{1}{2}\lambda_3$ and $\frac{1}{2}\lambda_8$. Thus, we need two quantum numbers to specify the states. $\frac{1}{2}\lambda_3$ is nothing but the isospin generator augmented by a zero in the (33) position, so we will still call it the strong isospin, with I_3 values of u, d and s being $+\frac{1}{2}, -\frac{1}{2}, 0$ respectively.

We define the second quantum number to be the *strangeness* S, defined as

$$S = \frac{1}{\sqrt{3}}\lambda_8 - \frac{1}{3}\,, \tag{87}$$

so that $S = 0$ for u and d and $S = -1$ for the strange quark [7]. Therefore, $S = +1$ for $\bar{s}$. A baryon is a combination of three quarks, so each quark carries a baryon number of $\frac{1}{3}$. We sometimes call $\lambda_8/\sqrt{3}$ the *strong hypercharge* (again, there is a weak counterpart, which is far more widely used) Y and write $Y = B + S$, but there is no more physics in this than (87).

- If we worked with only u and d quarks, that would have been an $SU(2)$ doublet of isospin, and so among the composites, I should have a triplet. Experimentally, the middle member of the triplet happens to be π^0. An $SU(2)$ isospin transformation can be performed by the interchange of u and d. There are two more $SU(2)$ doublets: K^0 and K^+, and their antiparticles, $\overline{K^0}$ and K^-.
- The roots of $SU(3)$ take one weight to the other. So they should act along the sides of the hexagon, and all members at the sides must belong to the same multiplet. By the isospin argument, π^0 should also belong to the same multiplet. So what are the irreducible representations: $\mathbf{7} \oplus \mathbf{2}$, or $\mathbf{7} \oplus \mathbf{1} \oplus \mathbf{1}$, or what? The answer is that for any $SU(n)$, $n \otimes \bar{n} = n^2 - 1 \oplus 1$. So the decomposition is $\mathbf{8} \oplus \mathbf{1}$.
- These nine lowest mass meson states can be divided into an octet and a singlet. There are three different sets of raising and lowering operators for the octet (they have no effect on the singlet state). One of them is the set of isospin raising or lowering operators, obtained from a combination of X_1 and X_2. The other two sets are called U-spin and V-spin respectively. V-spin operators are those obtained from X_4 and X_5, and can change an up quark to a strange quark and vice-versa. U-spin operators, obtained from X_6 and X_7, take a d quark to an s quark and back. Note that only U-spin does not change the charge of the quarks and hence of the mesons.

The lowest-mass mesons must have the orbital angular momentum $L = 0$, as higher L states, just like the hydrogen atom, contain the extra centrifugal energy $L(L+1)/r^2$. Mesons are $q\bar{q}$ composites, so can have a total spin S equal to 0 or 1. If we can write an inter-quark potential, the only nontrivial spin dependence can be of the form $\mathbf{s}_1.\mathbf{s}_2$ where $\mathbf{s}$ is the individual spin of the quark or antiquark. If $\mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2$, it is easy to show that in the singlet $S = 0$ state $\langle \mathbf{s}_1.\mathbf{s}_2 \rangle$ is negative, whereas in the triplet $S = 1$ state $\langle \mathbf{s}_1.\mathbf{s}_2 \rangle$ is positive. Thus, we expect the lowest mass mesons to have $S = 0$ and the next higher excitations, analogous to the hyperfine splitting in atoms, to have $S = 1$ (but still $L = 0$). The total angular momentum $J = L + S$ are 0 and 1 for these two cases.

For the $L = 0, S = 0$ combinations, there are nine states coming from $\mathbf{3} \otimes \bar{\mathbf{3}}$, divided into two irreducible representations of an octet and a singlet. Let us construct the easiest six of them, omitting the irrelevant minus sign:

$$
\begin{aligned}
|\pi^+\rangle &= |(I_3, S) = (1, 0)\rangle = u\bar{d}\,, \\
|\pi^-\rangle &= |(I_3, S) = (-1, 0)\rangle = d\bar{u}\,, \\
|K^+\rangle &= |(I_3, S) = (\frac{1}{2}, 1)\rangle = u\bar{s}\,, \\
|K^-\rangle &= |(I_3, S) = (-\frac{1}{2}, -1)\rangle = s\bar{u}\,, \\
|K^0\rangle &= |(I_3, S) = (-\frac{1}{2}, 1)\rangle = d\bar{s}\,,
\end{aligned}
$$

---

[7] This is an unfortunate convention, just like the one Benjamin Franklin used long ago, for which we are forced to show the electric current going in the opposite way of the flow of the current-carrying electrons.

$$|\overline{K}^0\rangle \;=\; |(I_3,S)=(\frac{1}{2},-1)\rangle = s\bar{d}\,. \tag{88}$$

The three pairs of ladder operators of SU(3), namely, $\lambda_1 \pm i\lambda_2$, $\lambda_4 \pm i\lambda_5$, and $\lambda_6 \pm i\lambda_7$, change one quark flavour into another and hence take one of these meson states to another. As all of them mix under SU(3), they must all be the part of the irreducible octet. In a two-dimensional I_3-S plot, these six points form the corner of a hexagon.

The other three states, formed out of $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$, all have $I_3 = S = 0$. All quantum numbers being the same, they can mix with one another to form the physical states. Among them, we identify π^0 as $(u\bar{u}-d\bar{d})/\sqrt{2}$, as isospin is still an almost exact symmetry. The singlet η_1 must contain all flavours in equal weightage so that it remains invariant under any SU(3) transformation, so we have

$$\eta_1 = \frac{1}{\sqrt{3}}\left(u\bar{u}+d\bar{d}+s\bar{s}\right)\,. \tag{89}$$

The last member of the octet, η_8, must be orthogonal to both π^0 and η_1, and can be found by the standard Gram-Schmidt orthogonalisation:

$$\eta_8 = \frac{1}{\sqrt{6}}\left(u\bar{u}+d\bar{d}-2s\bar{s}\right)\,. \tag{90}$$

The physical states η and η' are a mixture of η_1 and η_8, with η being dominantly the octet and η' dominantly the singlet. η' is heavier than η.

The $L=0, S=1$ mesons are a bit heavier than these lowest-lying mesons. The quark contents are almost identical, so they can be thought as the excited states of the ground-state mesons:

$$\begin{aligned}
|\rho^+\rangle &= |(I_3,S)=(1,0)\rangle = u\bar{d}\,,\\
|\rho^-\rangle &= |(I_3,S)=(-1,0)\rangle = d\bar{u}\,,\\
|K^{+*}\rangle &= |(I_3,S)=(\frac{1}{2},1)\rangle = u\bar{s}\,,\\
|K^{-*}\rangle &= |(I_3,S)=(-\frac{1}{2},-1)\rangle = s\bar{u}\,,\\
|K^{0*}\rangle &= |(I_3,S)=(-\frac{1}{2},1)\rangle = d\bar{s}\,,\\
|\overline{K}^{0*}\rangle &= |(I_3,S)=(\frac{1}{2},-1)\rangle = s\bar{d}\,.
\end{aligned} \tag{91}$$

The $I_3=0$, $S=0$ combinations are slightly different. The missing partner of the ρ-triplet is again $\rho^0 = (u\bar{u}-d\bar{d})/\sqrt{2}$, but the other two are different:

$$\omega = \frac{1}{\sqrt{2}}(u\bar{u}+d\bar{d})\,,\quad \phi = s\bar{s}\,. \tag{92}$$

We do not yet know why the mixing took place in such a manner as to make ϕ a pure $s\bar{s}$ bound state.

Only weak interaction violates S, so the lowest-mass K mesons, $K^\pm$ and $K^0,\overline{K}^0$, decay through weak interaction. The K^* mesons decay through $K^* \to K\pi$, which is a strong interaction. Similarly, all other vector mesons dominantly decay via strong interactions: $\rho,\omega\to\pi\pi$ and $\phi\to K^+K^-, K^0\overline{K}^0$. Why ϕ decays into two K mesons and not, say, to $\pi^+\pi^-\pi^0$, while the latter decay has more phase space? We will answer this puzzle soon.

Meson	J^P/J^{PC}	Mass (MeV)	Decay modes (dominant)	Type
$\pi^\pm$	0^-	139.57	$\mu^+\nu_\mu$	Weak
π^0	0^{-+}	134.98	$\gamma\gamma$	EM
η	0^{-+}	547.85	2γ (39.3%), $3\pi^0$ (32.6%) $\pi^+\pi^-\pi^0$ (22.7%)	EM, Strong
$K^\pm$	0^-	493.67	$\mu^+\nu_\mu$ (63.5%), $\pi^\pm\pi^0$ (20.7%)	Weak
$K^0, \overline{K}^0$	0^-	497.61	2π, 3π	Weak
η'	0^{-+}	957.66	$\pi^+\pi^-\eta$ (44.6%), $\pi^0\gamma$ (29.4 %) $\pi^0\pi^0\eta$ (20.7%)	Strong, EM
ρ	1^{--}	775.49	$\pi\pi$	Strong
ω	1^{--}	782.65	3π (89.2%), $\pi^0\gamma$ (8.9%)	Strong, EM
$K^{\pm *}$	1^-	891.66	$K\pi$	Strong
ϕ	1^{--}	1019.46	K^+K^- (49.2%), $K_L^0K_S^0$ (34.0%) $\pi^0\pi^+\pi^-$ (15.3%)	Strong

Table 1: The pseudoscalar and vector mesons. Only dominant decay modes are shown. Decays that involve γ are electromagnetic. K^{0*} mass and decay modes are similar to that of $K^{\pm *}$. K^0 and $\overline{K}^0$ are not mass eigenstates; they mix just like a two-level system and yield two mass eigenstates K_L^0, which dominantly decays to 3π, and K_S^0, which dominantly decays to 2π. This is the celebrated τ-θ puzzle.

Now, of course, we have come a long way from Gell-Mann's time, and we know of three more quarks. Charm (c) was discovered in 1974, and has a mass of about 1.4 GeV; bottom (b), discovered in 1985, is approximately 4.5 GeV; and the last one, top (t), discovered in 1996, is almost as heavy as a Gold nucleus — about 174 GeV! Charm and bottom hadronise to form mesons and baryons, but top decays before it can pick up some more quarks, and we never have a top hadron.

Can we extend the flavour symmetry group? Of course, we can, but with increasing meaninglessness. Introduction of charm enhances the group to $SU(4)$, which is a rank-3 group, and we need three quantum numbers to specify each state. The fundamental representation is four-dimensional, with four quarks, and the complex $\bar{\mathbf{4}}$ contains the antiquarks. We can again arrange the generators in such a way that the third quantum number is zero for u, d and s and $+1$ for c; this quantum number is also called *charm* and denoted by C, so don't confuse it with charge conjugation! Again, note that for charm, $C = +1$ for the quark, unlike the strange. Pretty unfortunate, and utterly confusing.

$SU(4)$ is badly broken, because charm is much heavier than the three lighter quarks. $SU(5)$, which includes b, is more badly broken; so there is no real meaning of talking about it, except that this brings another additive quantum number, B, where $B = -1$ for b, $+1$ for $\bar{b}$ and zero for all the rest. All these three quantum numbers, S, C, and B, are broken by weak interaction.

3.5 On the mass of the quarks

We can never see free quarks, so we do not know what their actual masses are. Rather, the quark mass appears as a parameter in the QCD Lagrangian: $\mathcal{L}_m = -m\bar{q}q$. Just like any other parameters of a renormalisable theory, m undergoes radiative corrections and runs, *i.e.*, the value of m depends on the scale Q where it is measured. When people talk about quark masses, they generally mean $m(Q)$, which is a parameter of the Lagrangian.

The perturbative running makes sense only in the regime where QCD is perturbative. The strong coupling constant α_s depend on the number of active flavours. To remind you, active flavours are those whose masses at a particular scale Q are less than Q. Thus, top is not an active flavour at $Q^2 = 10^4$ GeV^2. At a sufficiently low energy, of the order of a few hundred MeV, α_s becomes greater than unity and QCD becomes nonperturbative. The first three flavours, u, d, and s, are lighter than this scale, and that is why they are called light quarks. The other three are heavier, and called heavy quarks.

At the energy scale of the pion mass, we estimate $m_u \approx 5$ MeV, $m_d \approx 10$ MeV, and $m_s \approx 200$ MeV. How, then, is the pion so massive, and not of the order of 15 MeV? The answer lies beyond the scope of this course, but note that quark masses are not the only dimensionful parameters of the theory; there are two more, the pion decay constant f_π and the scale M where QCD becomes nonperturbative. The pion mass is a combination of all these. That the K mesons are heavier than the pions is mostly due to the mass difference of s and u/d, because f_π and f_K are close to each other. The point that should be stressed here is this: the degeneracy of proton and neutron, and the almost exact nature of strong isospin, is *not* a consequence of the almost degenerate nature of u and d; rather, it depends solely on the fact that the masses of these two quarks are much smaller than the QCD scale M — the exact values of the quarks masses are irrelevant.

3.6 P and C of mesons

The mesons are eigenstates of parity. Parity is a spatial operator and does not care about spin; so we expect the eigenvalue of parity to be $(-1)^L$, as happens for the hydrogen atom. However, there will be an extra minus sign; this comes from the relative minus sign between the intrinsic parities of quark and antiquark. The total parity is the product of the individual intrinsic parities and the orbital parity and so the eigenvalue is $-(-1)^L$. The lowest-lying mesons are $J = 0$ and negative parity; so they are called *pseudoscalar* mesons. Similarly, the next higher $J = 1$ and negative parity states are called *vector* mesons.

Not all mesons are eigenfunctions of the charge conjugation operator C; $\pi^0, \eta, \eta', \omega, \rho^0$ and ϕ are. Charge conjugation exchanges the particle with the antiparticle; both position and spin flip. Position flip is just parity and that gives $-(-1)^L$. If the meson is an $S = 1$ state, spin exchange gives a positive sign as the spin wavefunction is symmetric; if it is $S = 0$, we get a minus sign. Thus, the factor coming from spin exchange is $(-1)^{S+1}$. Multiplying, the eigenvalue of C is just $-(-1)^{L+S+1} = (-1)^{L+S}$. Thus, $C|\pi^0\rangle = |\pi^0\rangle$, $C|\phi\rangle = -|\phi\rangle$.

Are L and S good quantum numbers? In the Dirac theory, they are not, and there is no reason why they should be here too. However, if P is conserved, even and odd L states cannot mix. Similarly, if C is conserved, $S = 0$ and $S = 1$ states cannot mix. Of course, $L = 0$ and $L = 2$ states can mix.

3.7 Meson wavefunctions

The wavefunction is the product of the spatial and the spin wavefunctions (times the colour, but that is always singlet, so we will neglect that). For example, pions are spin-singlet, and hence

$$\begin{aligned}
|\pi^+\rangle &= \frac{1}{\sqrt{2}} u\bar{d}(\uparrow\downarrow - \downarrow\uparrow)\,, \\
|\pi^0\rangle &= \frac{1}{2}(u\bar{u} - d\bar{d})(\uparrow\downarrow - \downarrow\uparrow)\,.
\end{aligned} \tag{93}$$

Q. Write down the wavefunctions for the other pseudoscalar and vector mesons.

3.8 Meson decays and the Okubo-Zweig-Iizuka rule

A meson can decay in a number of ways. There are four, so to say, major categories.

- *Nonleptonic*: There are no leptons in the final state, only hadrons, the most common being a pair of mesons (or sometimes three). Let us concentrate on the decay to a pair of mesons $M_1 \to M_2 M_3$. A typical example is the decay of the $\overline{K}^0$ to $\pi^+\pi^-$. The quark level picture is $(s\bar{d}) \to (u\bar{d}) + (d\bar{u})$. Now there is a $\bar{d}$, the down antiquark, in both initial and final states. In a simplified picture, we can assume that this object was transferred unchanged from the initial state wavefunction to the final two-meson composite wavefunction. Such an object is called a *spectator quark* (note that whether it is a quark or an antiquark, we generally call it a quark). Leaving out the spectator, the quark-level decay is $s \to u\bar{u}d$. Of course, in the final state, u can combine with $\bar{u}$ and d with $\bar{d}$ and we can get two π^0 mesons, as an alternative possibility.

  How does the decay proceed? This is a weak interaction, because strangeness is not conserved (if it were, K mesons would have been stable objects). The s quark emits a virtual weak gauge boson (virtual, since the weak gauge bosons are way too heavy — almost 80 GeV — to be produced in a real state) and changes to an u. This changes charge, so the weak gauge boson must be a negatively charged one; we call it W^-. The same virtual W^- is emitted in the neutron β dcay. This W^- then breaks up into a d and an $\bar{u}$. Then the quarks and antiquarks fly apart, picking up a partner each, and hadronise; this is where strong interaction takes control. Unfortunately, there is no way to treat the hadronisation process from first principles.

  This is not the only possibility. The s can change to a d quark too, though the amplitude is highly suppressed [8]. This will eject out a neutral gauge boson, which can be the weak neutral Z, the photon, or even a gluon. This gauge boson then breaks up into an $u\bar{u}$ pair, and the same story of hadronisation continues.

  There are decays without a spectator, *e.g.*, $\phi \to \pi^+\pi^-\pi^0$. Here the initial s and $\bar{s}$ annihilate to produce a virtual intermediate state with total charge and strangeness zero, which then breaks up into a 3π final state. The dominant amplitude is the one where we have three gluons, radiated off the initial quarks, each of which going into either an $u\bar{u}$ or a $d\bar{d}$ pair. The decay width has g_s^{12} (each gluon gives g_s^2 to the amplitude) where g_s is the strong coupling constant, and hence this decay is highly suppressed even though it has a larger kinematic phase space compared to, say, $\phi \to K^+K^-$. For the latter decay, both the quarks are spectators, and a gluon radiated off either of them breaks into an $u\bar{u}$ pair; the amplitude goes only as g_s^2 and the decay width as g_s^4. Thus, processes with spectator quarks are more allowed than those with no spectators. This is known as the OZI (or Okubo-Zweig-Iizuka) rule.

- *Semileptonic*: Here the final state will contain two leptons and a meson, like $K^+ \to \pi^0\mu^+\nu_\mu$. The neutron β-decay, though it is for a baryon, is also a semileptonic process. The dynamics is almost identical with the nonleptonic case, but with marked simplification: the virtual W goes to a lepton pair, and leptons don't hadronise, so all the dirty strong interaction dynamics is absent.

---

[8] You will learn later that the leading amplitude for this process, $s \to d$, is zero in the Standard Model.

- *Radiative*: This includes at least one photon in the final state, like $B \to K^* \gamma$, which in the quark picture reads as $b \to s\gamma$. Again, the $b \to s$ amplitude is zero in the leading order, and the decay proceeds through higher-order quantum corrections — a one-loop effect. The decays $\pi^0 \to \gamma\gamma$ or $\eta \to \gamma\gamma$ are also radiative.
- *Leptonic*: This is a very rare type, where the quark-antiquark pair annihilate to form two leptons. However, this is theoretically very clean to handle. A good example is the decay $\pi^+ \to \mu^+ \nu_\mu$, which we will discuss in detail later. The only uncertainty comes from the fact that the meson is a bound system, so we cannot treat the initial state as a free quark-antiquark pair. As low-energy QCD — which is essentially nonperturbative — is at play, we cannot determine the difference between the bound state and the free $q\bar{q}$ pair from first principles. We hide our ignorance in so-called *decay constants*, which one must determine from the lifetime of the meson, or from simulating the meson and its decays on a discretized space-time, called a lattice.

Q. Identify the spectator in the following decays and write the quark-level decays: $\rho^+ \to \pi^+ \pi^0$, $B^0 \to J/\psi K^0$, $B_c \to J/\psi \pi^+$. Given, $B^0 = d\bar{b}$, $J/\psi = c\bar{c}$, $B_c = c\bar{b}$.

3.9 Baryons

There are 27 possible combinations from three triplets; however, due to the fermionic nature of barons, not all of them exist. First, note that there is a symmetric 10-plet (also called a decuplet), two octets, one being symmetric in its first two indices an antisymmetric in the third, and the other being antisymmetric in the first two, and a singlet.

Baryons are fermions (spin-$\frac{1}{2}$ or spin-$\frac{3}{2}$), so they obey Fermi statistics, so the total wavefunction must be antisymmetric under the exchange of any two constituent quarks. This brings in the concept of colour, but there are other reasons too to introduce a three-colour scheme for each quark.

- Consider the Δ^{++} baryon. This is one of the Δ resonances that are formed as an inermediate state in pion-nucleon scattering. For example,

  $$\pi^- + n \to \Delta^- \to \pi^- + n\,, \quad \pi^0 + n \to \Delta^0 \to \pi^0 + n, \pi^- + p\,, \quad \pi^+ + p \to \Delta^{++} \to \pi^+ + p\,. \tag{94}$$

  The Δ baryons, and their excited states, are formed through strong interaction and decay by the same; so their lifetimes are extremely short, of the order of 10^{-23} s, and they give a broad resonance. Obviously the quark content of Δ^{++} is uuu, and the space part of the wavefunction is manifestly symmetric. This is a spin-$\frac{3}{2}$ object; but if three spin-$\frac{1}{2}$ quarks combine to form a spin-$\frac{3}{2}$ state, the spin wavefunction must be completely symmetric too. However, one can write

  $$\psi = \psi_{\rm space} \psi_{\rm spin} \psi_{\rm colour}\,, \tag{95}$$

  so one can save the exclusion principle by postulating a colour part which is completely antisymmetric. Such a colour wavefunction can be written as

  $$\psi_{\rm colour} = \frac{1}{\sqrt{6}} \left[RGB + GBR + BRG - RBG - GRB - BGR \right]\,. \tag{96}$$

  Note the even and odd permutations. We postulate the colour wavefunction to be antisymmetric for all baryons, whether they are spin-$\frac{1}{2}$ or spin-$\frac{3}{2}$, and hence the product of space and spin wavefunctions *must be symmetric.* This we need to remember while constructing the baryon wavefunctions.

- Consider the process $e^+e^- \to f\bar{f}$, where f is any fermion, lepton or quark. This is an s-channel process and is allowed only if $\sqrt{s} \geq 2m_f$. This is known as the threshold for pair production of f. We can produce quark-antiquark pairs, but they will hadronise, and we have to sum over all the hadrons that are produced once we cross the threshold of a particular quark pair production. We define the ratio

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}\,. \tag{97}$$

  The ratio R can be measured very precisely. If we do not have any colour, *i.e*, there is only one possible pair of, say, $s\bar{s}$, the ratio will be totally different from the case where each s quark can come in three colours and the $s\bar{s}$ production cross-section is multiplied by a factor of 3 (it is close to 3.14, taking the quantum corrections into account). The experiments decisively show that there are three colours.

- Consider the decay $\pi^0 \to \gamma\gamma$. This is a one-loop process. The amplitude, which is a three-point Green's function, has three legs connected to the π^0 and two photons respectively. The particle going inside the loop is a quark. As this can come in three colours, we have to multiply the final result by three, and the data is completely consistent with the three-colour scheme.

Let us first concentrate on the proton uud. This belongs to the $\mathbf{8}$ of $\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3}$ of SU(3), but which $\mathbf{8}$? The answer is that it is a superposition from both $\mathbf{8}$s. Remember the Young tableaux: the first one is symmetric in its first two quarks, and antisymmetric in the third, while the second one is antisymmetric in its first two quarks. Thus, we can write, for the flavour part,

$$p_S = \frac{1}{\sqrt{6}}\left[udu + duu - 2uud\right]\,, \quad p_A = \frac{1}{\sqrt{2}}\left[udu - duu\right]\,. \tag{98}$$

The spin wavefunctions must behave identically, as the product is going to be completely symmetric. Thus,

$$\chi_S = \frac{1}{\sqrt{6}}\left[\uparrow\downarrow\uparrow + \downarrow\uparrow\uparrow - 2\uparrow\uparrow\downarrow\right]\,, \quad \chi_A = \frac{1}{\sqrt{2}}\left[\uparrow\downarrow\uparrow - \downarrow\uparrow\uparrow\right]\,. \tag{99}$$

The total wavefunction, where the proton has a spin-projection $+\frac{1}{2}$, is given by

$$\frac{1}{\sqrt{2}}\left[p_S\chi_S + p_A\chi_A\right]\,, \tag{100}$$

which, under a brute-force decomposition, gives

$$|p\uparrow\rangle = \frac{1}{\sqrt{18}}|u\uparrow u\downarrow d\uparrow + u\downarrow d\uparrow u\uparrow - 2u\uparrow u\uparrow d\downarrow + \text{permutations}\rangle\,. \tag{101}$$

For neutron, it is enough to replace u and d by d and u respectively.

3.10 Magnetic moment of baryons

The magnetic moment of a point charge Qe and mass m is defined as $\mu = Qe/2m$. The magnetic moment of the proton (which, by convention, is defined in the highest possible spin projection state) can be written, in terms of the constituent quarks, as

$$\begin{aligned}\mu_p &= \sum_{i=1}^{3}\langle p\uparrow|\mu_i(\sigma_3)_i|p\uparrow\rangle \\ &= \frac{1}{18}\left[(\mu_u-\mu_u+\mu_d)+(-\mu_u+\mu_u+\mu_d)+4(2\mu_u-\mu_d)\right]\times 3 \\ &= \frac{1}{3}(4\mu_u-\mu_d)\,. \end{aligned} \tag{102}$$

Q. Derive an analogous expression for the magnetic moments of neutron.

Q. If $m_u = m_d$, show that

$$\frac{\mu_n}{\mu_p} = -\frac{2}{3}\,. \tag{103}$$

Compare it with the experimental value of -0.68498. What value would you expect if both proton and neutron were point particles?

Q. What is the magnetic moment of a Δ^{++} baryon?

3.11 Colour

That each quark comes in three colours is also borne out by experiments. Consider the process $e^+e^- \to q\bar{q}$. Its amplitude is simply related to the amplitude of the QED process $e^+e^- \to \mu^+\mu^-$ by the replacement of muon charge of -1 by the quark charge e_q. Thus, if the energy is such that all the light quarks u, d, and s can be produced, but nothing else, we have

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = N_c\left[e_u^2+e_d^2+e_s^2\right] = \frac{2}{3}N_c\,, \tag{104}$$

where we have used the fact that all light quarks produced end up in hadrons. This ratio comes out to be close to 2, so there are three colours. If the energy is increased so that $c\bar{c}$ pairs can also be produced (this is known as crossing the charm threshold), R jumps up to 10/3, as expected from the extra e_c^2 term. With even higher energies above the b threshold, R jumps further to 11/3, which vindicates the colour hypothesis.

4 Weak Interaction

Weak interaction causes neutron β-decay, and the decays of the K mesons, muons, and charged pions. It is a very short-range force; the effect does not go beyond the dimension of a typical nucleus. The fact that it is weak is reflected in the typical lifetimes of these decays; they are of the order of 10^{-6}-10^{-8} second, much larger than those mediated by strong ($\sim 10^{-23}$) or electromagnetic ($\sim 10^{-15}$) second.

Weak interaction is special in the sense that it violates a number of symmetries. The first one is the strangeness-like set: S, C, and B. Note that there is no such type-violation for leptons; one does not see the violation of electron or muon numbers, for example, the decay $\mu^- \to e^-e^+e^-$ is never seen. What causes this difference? As we will learn later, this is due to the approximately massless nature of the neutrinos. The individual lepton number conservations would have been exact if the neutrinos are all massless (or have the same mass). This is true to a very good approximation, so we do not see electron or muon number violations. However, the neutrinos are not exactly massless, and later we will see how the conversion of one neutrino species to another takes place, violating the individual lepton numbers. Thus, $\mu^- \to e^-e^+e^-$ does take place, but at an unobservably small rate.

The second thing that is violated is parity. Weak interaction treats left and right differently. This is most obvious from the property of neutrinos, because they can feel only the weak force (why?). However, this does not mean that all weak interactions should involve neutrinos; for example, the decay $K^0 \to \pi^+\pi^-$ does not. But the presence of neutrinos is an irrefutable sign of weak interaction. We observe only left-chiral neutrinos in nature. If neutrinos are massless, chirality and helicity are identical, so their helicity is also left: the projection of spin on the momentum axis is $-\frac{1}{2}$. That there is a left-right asymmetry was first shown by Wu and her collaborators; that the neutrino is totally left-handed was demonstrated in another famous experiment by Goldhaber et al.

Weak interaction also violates charge conjugation. One observes $\pi^+ \to \mu^+\nu_{\mu L}$ but not $\pi^- \to \mu^-\nu_{\bar{\mu}L}$; in fact, there is no such object, as far as we know, as $\bar{\nu}_L$. However, the decay rates of $\pi^+ \to \mu^+\nu_{\mu L}$ and $\pi^- \to \mu^-\nu_{\bar{\mu}R}$ are identical to the experimental precision. They are related by the combined operation CP (or PC, the order is irrelevant); P changes the chirality and C changes the particle to the antiparticle. Thus, CP is apparently conserved in weak interaction. From the CPT theorem, this implies that weak interaction is invariant under time-reversal too.

4.1 Fermi's theory of allowed β-decay: nonrelativistic treatment

We will go through a nonrelativistic derivation of Fermi's theory of β decay. From a relativistic point of view, the derivation is not very rigorous; however, the energy associated with β decay is of the order of MeV, where all nucleons and nuclei are very much nonrelativistic. The electron and the antineutrino are not so; thus, they should be treated with Dirac or Weyl fermions. While the derivation is not unduly complicated, it does not shed any more light than the nonrelativistic treatment, which we shall follow. Here, we will take usual one-component wavefunctions for both the leptons.

The decay follows an exponential decay pattern, typical for radioactive decays, and the number of parent particles goes down as $N(t) = N(0)\exp(-\lambda t)$, where $\tau_{\rm mean} = 1/\lambda$ is the mean life and $\tau_{1/2} = \ln 2/\lambda = 0.693\tau_{\rm mean}$ is the half-life. The decay involves a transition from an initial nuclear state $|i\rangle$ to a final state $|f\rangle$, and Fermi's Golden rule gives the transition probability,

$$\lambda = 2\pi |H_{fi}|^2 \rho(E_0)\,. \tag{105}$$

Note that we have used the natural units with $\hbar = 1$, otherwise the first factor would have been $2\pi/\hbar$. The transition matrix element H_{fi} is

$$H_{fi} = \langle Y, e^-, \bar{\nu}_e | \mathcal{H}_{\rm weak} | X\rangle\,, \tag{106}$$

where $\mathcal{H}_{\rm weak}$ is the weak Hamiltonian responsible for the β decay. E_0 is the total energy available in the final state, *i.e.*,

$$E_0 = Q_{\beta^-} + m_e + m_\nu = M_n(A,Z) - M_n(A,Z+1)\,, \tag{107}$$

and $\rho(E_0)$ is the phase space factor for the final state leptons.

We take $\mathcal{H}_{\rm weak}$ to be a point interaction, *i.e.*, all fermions interacting at the same space-time point. In other words, we have two currents; a leptonic current of the form $\bar\psi_e\Gamma\psi_{\bar\nu}$ (so that an electron and an antineutrino are created), and a hadronic current of the form $\bar\psi_p\Gamma\psi_n$. Γ is some combination of the γ-matrices; the original Fermi theory used $\Gamma = \gamma^\mu$, so in momentum space, the interaction looked like

$$\mathcal{L}_{\rm weak} = G_F(\bar e\gamma^\mu\nu)(\bar p\gamma_\mu n)\,. \tag{108}$$

Here, the particle names stand for the momentum-space spinors, not the operators. The exponential factors have been taken care of by the δ-function $\delta(x_h - x_l)$ and an integration over x_h (or x_l), where x_h and x_l are the four-coordinates of the hadronic and leptonic vertices respectively, and the δ-function is just a mathematical way of saying that it is a contact interaction.

There are two points that you should note right here. First, the theory cannot be parity-violating; we know from our experiences in quantum electrodynamics that any theory based on vector bilinear currents must be invariant under parity, just like QED. In fact, with the knowledge that weak interaction occurs only for left-chiral fermions (for the experiment, read on), it is easy to modify the form:

$$\mathcal{L}_{\rm weak} = \frac{G_F}{\sqrt{2}}(\bar e\gamma^\mu(1-\gamma_5)\nu)(\bar p\gamma_\mu(1-\gamma_5)n)\,, \tag{109}$$

where the factor of $\sqrt{2}$ is just for normalization. We will, however, neglect the parity violation aspect in our subsequent nonrelativistic discussion. The second point is more important. The mass dimension of four fermion fields is M^6, but the Hamiltonian density should have a dimension of M^4, so $[G_F] = M^{-2}$. But a theory with coupling whose dimension is of any negative power of mass is bound to be nonrenormalizable. In other words, it is meaningless to calculate radiative corrections to such a theory; it is meaningless even to extrapolate it to high energies. For example, if you want to know what the weak interaction contributions are for the process $e^+e^- \to \mu^+\mu^-$ at $\sqrt{s} = 100$ GeV, you cannot find it out with the Fermi theory! But Fermi theory was not formulated for that; it does very well what it was supposed to do, to calculate the β-spectrum at low energies (a few MeV at the most).

This, then, is a good point to introduce the idea of *effective theories*. An effective theory is what its name suggests; a theory that is valid up to a certain energy scale, and not above that. When you go above that scale, you have to use a more complete, and hopefully better, theory. For Fermi theory, the role is played by the gauge theory of weak interaction. This is also known as the *ultraviolet completion* (at higher energies, hence ultraviolet) of the theory. As you reach the threshold from below, the effective theory gradually ceases to be a good description. You can, of course, work with the full theory at any energy scale; but you will find that whether you use the complete weak interaction gauge theory or Fermi theory to tackle β-decay, the results are same, and the latter is much easier to use.

You can now immediately see why the Fermi coupling G_F has a mass dimension M^{-2}. Consider the β-decay amplitude in the full theory; we can also write it as $n + \nu_e \to p + e$, which is diagrammatically the same as the QED $2 \to 2$ scattering processes. Thus, the amplitude should be proportional to the hadronic current, times the leptonic current, times the gauge coupling g squared (occuring at two vertices), times the weak gauge boson propagator, and some numerical factors. But the weak gauge bosons are not massless like the photon; so we have to write the corresponding propagator as $1/k^2 - M_W^2$, where k is the momentum transfer (and for β-decay, at most of the order of a few MeV, much smaller than the W-boson mass of 80.41 GeV). So, apart from numerical factors,

$$G_F \propto \frac{g^2}{k^2 - M_W^2} \approx \frac{g^2}{M_W^2}\,, \tag{110}$$

where we have used the fact that $k^2 \ll M_W^2$. Obviously, the theory starts to break down as k^2 approaches M_W^2.

Now let us come back to the nonrelativistic treatment again.

The Schrödinger equation does not involve spin, so we can work with single-component wavefunctions for the fermions. We can write

$$H_{fi} = G_F \int \psi_f^* \psi_i \phi_e^* \phi_{\bar{\nu}}^* \; d^3\mathbf{r}\,, \tag{111}$$

where i and f stand for initial and final state nucleons respectively. The integration is over the nuclear volume.

Let us assume the lepton wavefunctions to be plane waves. This is a good approximation for the neutrino which interacts very weakly, but may not be a good approximation for the electron, which will be affected by the Coulomb field of the nucleus. The effect will be more prominent for high-Z nuclei; also, there will be repulsion for positrons and attraction for electrons. We will deal with this Coulomb correction later. Right now, assume the plane waves to be normalized over a volume V:

$$\int \phi_e^* \phi_e \; d^3\mathbf{r} = \frac{1}{V}\,. \tag{112}$$

So,

$$\phi_e = \frac{1}{\sqrt{V}} \exp(i\mathbf{k}_e.\mathbf{r})\,, \qquad \phi_{\bar{\nu}} = \frac{1}{\sqrt{V}} \exp(i\mathbf{k}_\nu.\mathbf{r})\,, \tag{113}$$

and the leptonic part of the matrix element is

$$\frac{1}{V} \exp(-i(\mathbf{k}_e + \mathbf{k}_\nu).\mathbf{r}) = \frac{1}{V} \exp(-i\mathbf{k}.\mathbf{r})\,, \tag{114}$$

where $\mathbf{k} = \mathbf{k}_e + \mathbf{k}_\nu$ is the momentum transferred to the leptons. However, $|\mathbf{k}| \sim 1$ MeV, $|\mathbf{r}| \sim$ 1-5 fm, so $|\mathbf{k}.\mathbf{r}| \ll 1$ (remember that in the natural system, $1 = \hbar c = 197$ MeV-fm). Thus, one can write

$$\exp(-i\mathbf{k}.\mathbf{r}) = (1 - i\mathbf{k}.\mathbf{r} + \cdots) \approx 1\,. \tag{115}$$

We can also do a multipole expansion of the exponential. The first term, which is for $L = 0$, corresponds to the dominant transitions, called *allowed*, and $L \geq 1$ terms correspond to suppressed transitions, called *forbidden.* Note that forbidden transitions do occur but with less probability and hence with greater half-life.

We will consider only allowed transitions, so

$$H_{fi} = \frac{G_F}{V} \int \psi_f^* \mathcal{O} \psi_i \; d^3\mathbf{r} = \frac{G_F}{V} \langle f|\mathcal{O}|i\rangle \; = \frac{G_F}{V} M_{ij},, \tag{116}$$

where $\mathcal{O}$ is the operator that causes the $n \to p$ transition. Unfortunately, the matrix element M_{ij} is not known; the initial and final states are nuclear states, *i.e.*, bound states of several nucleons, and there is no way one can calculate the matrix element from first principles. However, if $|i\rangle$ and $|f\rangle$ are same shell-model states, $M_{ij} = 1$; they are called *super-allowed* transitions.

The next task is to calculate $\rho(E_0)$.

There are two leptons in the final state; so the phase space is 12-dimensional (6 for position coordinates and 6 for momenta) [9]. We know that volume of unit cell in a $2N$ dimensional phase space is $(2\pi)^N$; in conventional units, it is just h^N. Number of states in a particular volume element is given by the product of number of electron states and number of antineutrino states, times a δ-function to conserve energy-momentum. You have already learnt this earlier, so we can just write down the result:

$$d\rho = \frac{1}{(2\pi)^6}\left(p_e^2 dp_e d\Omega_e\right)\left(p_\nu^2 dp_\nu d\Omega_\nu\right) d^3\mathbf{r}_e d^3\mathbf{r}_\nu \delta(E_0 - E_e - E_\nu)\,. \tag{117}$$

E_e and E_ν are the total energies of the leptons and from eq. (107), the δ-function is nothing but energy conservation. Each angular integration gives 4π, and each spatial volume integration gives a factor of V. Also, $E^2 = p^2 + m^2$, so $EdE = pdp$, where p is the magnitude of the three-momentum. For the antineutrino, we can neglect its mass and write $p_\nu = E_\nu$. Thus

$$d\rho = \frac{V^2}{4\pi^4}(p_e^2 dp_e)(E_\nu^2 dE_\nu)\delta(E_0 - E_e - E_\nu)\,. \tag{118}$$

The antineutrino is not observed, so we integrate over dE_ν using the δ-function and get

$$d\rho = \frac{V^2}{4\pi^4} p_e^2 (E_0 - E_e)^2 dp_e\,. \tag{119}$$

Putting all this together, we get the decay constant for any particular electron three-momentum p_e:

$$d\lambda(p_e) = \frac{1}{2\pi^3} G_F^2 |M_{fi}|^2 p_e^2 (E_0 - E_e)^2 dp_e\,, \tag{120}$$

which is expectedly independent of V (so we could have done unit volume normalization from the very beginning). Also convince yourself, by dimensional analysis or otherwise, that in the conventional system the expression is to be divided by $c^3\hbar^7$.

Eq. (120) neglects an important correction, namely, the fact that the outgoing electron (or positron) wavefunction is not a plane wave but is modified by the Coulomb field of the nucleus. The multiplicative correction factor is given by

$$F(Z, E_e) \equiv \frac{|\phi_e(0)|^2}{|\phi_e(0)|^2_{\text{free}}} = \frac{2\pi\eta}{1 - e^{-2\pi\eta}}\,, \tag{121}$$

where $\eta = Ze^2/v$ for electrons and $-Ze^2/v$ for positrons, v being the asymptotic velocity of the β-particle, *i.e.*, the velocity far away from the nucleus. (Try to convince yourself that in the SI system, η should be $\pm Ze^2/\hbar v$.) For small Z (low-Z nuclei) or large v (more energetic β-decay), $\eta \to 0$ and $F \to 1$.

Thus, incorporating the Coulomb correction,

$$d\lambda(p_e) = \frac{1}{2\pi^3} G_F^2 |M_{fi}|^2 p_e^2 (E_0 - E_e)^2 F(Z, E_e) dp_e\,. \tag{122}$$

If the number of β-particles emitted per second in the momentum interval p_e and $p_e + dp_e$ be $dN(p_e) = N(p_e)dp_e$, then

$$dN(p_e) = N_0 d\lambda(p_e)\,. \tag{123}$$

---

[9] Actually, there are three particles in the final state, but we take the phase space for proton to be unity. Anyway, it is highly nonrelativistic, and we do not detect it, so we should integrate over its phase space. The remnant is in the energy-conserving δ-function.

This tells us that for a given momentum p_e, the expression $\sqrt{N(p_e)/p_e^2F(Z,p_e)}$ plotted against $E_0 - E_e$ will give a straight line, and the maximum energy of the β-particles is where $E_e = E_0$. This straight line is known as the *Kurie plot.*

A common practice is to express everything in terms of the electron rest mass:

$$W = \frac{E_e}{m_e}\,, \qquad W_0 = \frac{E_0}{m_e}\,, \qquad \bar{p} = \frac{p_e}{m_e}\,. \tag{124}$$

So

$$d\lambda(\bar{p}) = \frac{1}{2\pi^3}G_F^2|M_{fi}|^2 m_e^5\bar{p}^2F(Z,\bar{p})(W_0 - W)^2 d\bar{p}\,, \tag{125}$$

and

$$\lambda(\bar{p}) = \frac{1}{2\pi^3}G_F^2|M_{fi}|^2 m_e^5 f(Z,p_0)\,, \tag{126}$$

$$f(Z,p_0) = \int_0^{p_0} \bar{p}^2F(Z,\bar{p})(W_0 - W)^2 d\bar{p}\,. \tag{127}$$

For allowed transitions, only the first term in the plane wave expansion is retained, so that corresponds to $L = 0$. As this is the total orbital angular momentum of the two leptons, this must be the change in orbital angular momentum of the nucleus. Thus, the parity of the nucleus does not change in an allowed transition. The two leptons coming out of the nucleus can have a total spin of 0 or 1. Suppose they are in the total spin $S = 0$ combination. Then the total spin change of the nucleus $\Delta J = 0$. If the leptons come out in $S = 1$ state, the total angular momentum carried by the leptons is one unit too, so $\Delta J = 0$ or ± 1 for the nucleus, except that $0 \leftrightarrow 0$ transition is forbidden in this case as angular momenta 0 and 1 in the final state cannot make up an angular momentum 0 in the initial state. The first type of transitions ($S = 0$) are called *Fermi processes*, and the second type ($S = 1$) are called *Gamow-Teller processes.* Note that there can be mixed transitions too.

For the first forbidden transitions, $L = 1$, so the parity of the nucleus changes. For Fermi processes $\Delta J = 0, \pm 1$ (this comes from the orbital angular momentum), but $0 \to 0$ is forbidden. For Gamow-Teller transitions, $S = 1$, and $\Delta J = 0, \pm 1, \pm 2$.

4.2 Wu's experiment

Wu and her collaborators worked with radioactive Cobalt-60. Co^{60} has a spin-parity of $J^P = 5^+$ (in nuclear physics terminology, spin means the total angular momentum of the nucleus). It decays to the excited state of Ni^{60}, which has a spin-parity of 4^+, through a β-transition. The parity does not change, and the spin changes by one unit, so it is a Gamow-Teller transition. The Co^{60} nucleus de-excites to its ground state 0^+ emitting a couple of γ-ray photons, of energies 1.173 and 1.332 MeV.

The sample was placed in a magnetic field. Cobalt being ferromagnetic, the nuclear spins should be all aligned. This polarisation was achieved using a technique called *adiabatic demagnetisation.* The alignment information is retained in Ni^{60}, so the rates of γ emission in the polar and equatorial directions with respect to the applied magnetic field would be different.

When Co^{60} decays to Ni^{60} through

$$Co^{60} \to Ni^{60*} + e^- + \bar{\nu}_e\,, \tag{128}$$

The leptons carry the one unit of angular momentum. This, then, must be aligned along the initial direction of spin polarisation. If weak interaction respected parity, we would have expected equal amount of $e_L\bar{\nu}_R$ and $e_R\bar{\nu}_L$. For the first, the antineutrino has positive helicity, so goes along the direction of spin, i.e., the direction of polarisation. The electron then must go against the polarisation. For the second case, the direction of the leptons would be reversed. If the number of electrons towards and against the direction of polarisation be different, that tells us that parity is not respected in weak interaction.

Wu et al. found that the numbers are different; the asymmetry flips if the magnetic field is flipped, and vanishes if the sample is warmed (which kills the polarisation), so that the effect is due to spin polarisation and not due to the external magnetic field.

After this groundbreaking experiment, several other experiments also found the signature of parity violation in the decay of $\pi^+ \to \mu^+\nu_\mu$, relying on the fact that if weak interaction violates parity, the muon will develop a polarisation in flight.

Wu's experiment forced a parity-violating component in the weak interaction Lagrangian. The original Lagrangian suggested by Fermi, (108), is parity conserving, and the easiest way is to replace γ^μ by $\gamma^\mu(1 - a\gamma_5)$. Is a close to 1 or small? If $a = 1$, the parity violation is maximal, the weak interaction couples only to the left-chiral fermions (as $1 - \gamma_5$ projects out the left-chiral component only). If $a \neq 1$, one can write

$$1 - a\gamma_5 = \frac{1-a}{2}(1+\gamma_5) + \frac{1+a}{2}(1-\gamma_5)\,,$$

so that both left and right-chiral fermions feel the weak force. The next step is to show that a is indeed unity.

4.3 Goldhaber experiment

In 1957, Goldhaber, Grodzins, and Sunyar performed a classic experiment to show conclusively that neutrinos are entirely left-handed. In other words, they showed that the charged current weak interaction, which was earlier demonstrated by Wu and her collaborators to have a Lorentz structure of $\gamma^\mu(1 - a\gamma_5)$ — note that as long as $a \neq 0$ this will violate parity — violates parity maximally, i.e., $a = 1$. This is how the experiment went.

The radioactive isotope Europium-152 has total angular momentum $J = 0$. This captures one of its own electron and transmutes to Samarium-152, emitting a neutrino:

$$Eu^{152} + e \to Sm^{152*} + \nu\,, \tag{129}$$

where the asterisk indicates that Sm^{152} is in an excited state. Obviously, the neutrino and the Samarium nucleus should go in the opposite direction, conserving the linear momentum. The excited Samarium has a total angular momentum $J = 1$. The only spin that entered the picture is the spin of the captured electron, which is $\frac{1}{2}$, so the neutrino, with $J = \frac{1}{2}$, and Sm^{152*}, with $J = 1$, must have their spins aligned in the opposite direction, so that they add up to $\frac{1}{2}$. If the neutrino is right-handed, its spin is along is velocity, so is that of the Samarium nucleus. If the neutrino is left-handed, the spin is opposite to its velocity, and Sm^{152*} follows that pattern too. Excited Samarium de-excites into the ground state of Sm^{152} by emitting a 960 keV photon. The lifetime is only about 4×10^{-14} sec, so the Samarium nucleus has no chance to come into a thermal equilibrium, even if it is buried in a Europium sample, and the information on its direction of motion is retained by the photon. The photon, with $J = 1$, also carries the spin information of Sm^{152*}. Thus, *the spin of such an elusive particle as the neutrino can be measured by measuring the polarisation of the photon!*

Suppose the neutrino is left-handed. Then Sm^{152*} will have the momentum and the spin aligned in the opposite direction. If we catch those photons that move in the forward direction, i.e., the direction of motion of the Samarium nucleus, they will also have the spin against the direction of motion; these are left-handed photons. Similarly, if the neutrino is right-handed, the forward-moving photons will be right-handed, too. In short, the spin of the forward moving photons will be the same as that of the neutrino.

How to know which photons are forward moving? For this, we use resonant scattering: $\gamma + Sm^{152} \to Sm^{152*} \to Sm^{152} + \gamma$. The idea is this. While the level-splitting between the ground state and the excited state of Samarium is 960 keV, the energy of the emitted γ-ray is not exactly that, because some energy is taken up by the recoiling Samarium. If these less energetic γ rays fall upon ground-state Samarium, they cannot excite them, because not enough energy is available. Only those photons moving in the forward direction can cause the resonant scattering; they are slightly more energetic than 960 keV from the forward kick of the parent excited Samarium, and this extra energy allows for some recoil in the resonant scattering.

Magnetised iron can act as a filter; photons polarised antiparallel to the spin polarisation of the electrons of iron are more absorbed than photons polarised parallel. These photons can be detected with a sodium iodide scintillator attached to a photomultiplier tube. Goldhaber *et al.* found that all photons that can cause resonant scattering are left-handed, and that shows conclusively that all neutrinos are left-handed too.

4.4 Lessons from the Goldhaber experiment

- The neutrinos are completely left-handed, or left-chiral (they are the same for massless particles). But the vector current $\bar\psi\gamma^\mu\psi$ treats both L and R fields equally, so the form of the current should be modified to take parity violation into account.

- Consider the lepton current. We can write it as

$$\bar e\gamma^\mu\nu_L = \bar e\gamma^\mu P_L\nu = \bar e P_R\gamma^\mu P_L\nu\,, \tag{130}$$

  so that only the L-fields see the weak interaction and not the R-fields. It is expected then that the same will be true for the quark current, and the Fermi Lagrangian (108) gets modified to (109).

- Weak interaction changes electric charge, so we need charged carriers. We have also seen the neutral current processes, like the neutrino scattering off nucleons $\nu N \to \nu N$, which must be weak interaction because neutrinos are involved, so we need neutral carriers too. These three objects, also known as *weak gauge bosons*, will be called W^+, W^-, and Z respectively. This is to be compared with the single charge-neutral photon of electrodynamics.

- Weak interaction is short-range, of the order of a fermi or even less, so weak gauge bosons must be massive (remember that 0.2 GeV-fm ≈ 1).

- The gauge group of electromagnetism is U(1). This has only one generator and we have one gauge boson in the theory. This is a general feature — the number of gauge bosons is equal to the number of generators of the group (can you argue why?). For weak interaction, we need a more complicated group. SU(2) will do very well, this has three generators.

- The photon mass term $-m^2 A_\mu A^\mu$ is not gauge invariant, but photon is massless, so there is no problem in electromagnetism. However, there seems to be no way to give masses to the weak gauge bosons in a gauge-invariant way.
- The L-fermions take part in the weak interaction, so they must transform non-trivially under weak SU(2). The simplest representation is the fundamental one, the doublet **2**. The R-fields do not see weak interaction, so they must be singlets.
- The mass term in the Dirac Lagrangian, $-m\bar{\psi}\psi = -m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$, is bound to be a doublet, and so cannot be there in the Lagrangian.
- Bottomline: if the weak interaction gauge group is SU(2), or anything higher, there is no way to give masses to the gauge bosons and fermions in a gauge-invariant way. All of them should remain massless. We will soon see how this problem is overcome.

4.5 Digression: kinematics of scattering and decay

The easiest process to deal with is the two-body decay $A \to B + C$. We take the particle A to be at rest; if it is not, we have to boost back all the quantities to the lab frame where A is moving.

If A is at rest, B and C must go at the opposite direction with equal momenta. Let the four-momenta of A, B and C be p_A, p_B and p_C respectively. We can write

$$p_A = (M, \mathbf{0}), \quad p_B = (E_B, \mathbf{p}), \quad p_C = (E_C, -\mathbf{p}). \tag{131}$$

Obviously, $M = E_B + E_C$, so we can write, following the relativistic energy-momentum relationship,

$$M = \sqrt{m_B^2 + p^2} + \sqrt{m_C^2 + p^2}, \tag{132}$$

where $p = |\mathbf{p}|$. After some easy algebraic manipulation, we obtain

$$\begin{aligned} p &= \frac{1}{2M}\sqrt{M^4 + m_B^4 + m_C^4 - 2M^2 m_B^2 - 2M^2 m_C^2 - 2m_B^2 m_C^2} \\ &= \frac{1}{2M}\left[\{M^2 - (m_B - m_C)^2\}\{M^2 - (m_B + m_C)^2\}\right]^{\frac{1}{2}}. \end{aligned} \tag{133}$$

This immediately tells us that the decay is not allowed if $M < m_B + m_C$. Also, if $m_B = m_C = m$ (which happens is B and C are either identical particles or antiparticles of one another),

$$p = \frac{1}{2}\sqrt{M^2 - 4m^2}. \tag{134}$$

Q. The neutrinos coming from the sun mostly originate from the reaction $p + p \to d + e^+ + \nu_e$. There are some electrons coming from the reaction $Be^7 + e^- \to Li^7 + \nu_e$. The former neutrinos show an energy distribution, the latter neutrinos have a definite energy. Why?

4.5.1 $2 \to 2$ scattering

It is always a good policy to express the observables in terms of the quantities that are Lorentz invariant; remember that neither energy nor three-momentum are Lorentz-invariant quantities.

Consider the process $A + B \to C + D$. Given the masses, there are only two independent quantities for this scattering: the centre-of-mass (CM) energy and the scattering angle.

While it is advisable to calculate everything in the CM frame, this may not be identical with the lab frame (*e.g.*, in a fixed-target experiment where the beam of B hits the fixed target A). Let us take the beamline, *i.e.*, the direction of motion of A and B, to be the z direction. Let the direction of C be in the y-z plane, making an angle θ with the z axis (this is known as the scattering angle). To have momentum balance, D must also lie in the y-z plane, and in the CM frame, must go opposite C. Thus we can write

$$\begin{aligned} p_A &= (E_A, 0, 0, p)\,, \\ p_B &= (E_B, 0, 0, -p)\,, \\ p_C &= (E_C, 0, p\sin\theta, p\cos\theta)\,, \\ p_D &= (E_D, 0, -p\sin\theta, -p\cos\theta)\,. \end{aligned} \tag{135}$$

4.5.2 Mandelstam variables

The scattering can most conveniently be expressed in terms of three Lorentz scalars, namely,

$$s = (p_A + p_B)^2, \quad t = (p_A - p_C)^2, \quad u = (p_A - p_D)^2, \tag{136}$$

where we have followed the convention that p_A and p_B are incoming momenta and p_C and p_D are outgoing momenta; also, $p^2 = p_\mu p^\mu$. The variables s, t and u are known as *Mandelstam variables.* Note that with four-momentum conservation, we can also write

$$s = (p_C + p_D)^2, \quad t = (p_B - p_D)^2, \quad u = (p_B - p_C)^2, \tag{137}$$

but we will follow the first convention (136). All the particles are real and hence $p_i^2 = m_i^2$, this helps us to express any four-dimensional dot product in terms of the Mandelstam variables and masses. For example,

$$p_A.p_B = \frac{1}{2}\left(s - m_A^2 - m_B^2\right)\,. \tag{138}$$

The variables are not all independent:

$$\begin{aligned} s + t + u &= (p_A + p_B)^2 + (p_A - p_C)^2 + (p_A - p_D)^2 \\ &= p_A^2 + p_B^2 + p_C^2 + p_D^2 + 2p_A.(p_A + p_B - p_C - p_D) = \sum_{i=A}^{D} m_i^2\,, \end{aligned} \tag{139}$$

where we have used the momentum conservation $p_A + p_B - p_C - p_D = 0$. With two independent variables, we can replace p and θ with the Mandelstam variables.

Q. Express $p_A.p_C$, $p_B.p_C$, $p_A.p_D$, $p_B.p_D$ and $p_C.p_D$ in terms of the Mandelstam variables and masses.
Q. In the ultra-relativistic limit, we can neglect the rest masses of the particles compared to their total energy. Show that in this limit $s = 4E^2$, $t = -2E^2(1 - \cos\theta)$, $u = -2E^2(1 + \cos\theta)$.

4.5.3 Cross-section

We would like to know how many events we are going to observe in a scattering process. If the particles retain their identity, it is called an elastic scattering; otherwise, it is inelastic. First, we would like to define something intrinsic; some sort of probability that A will interact with B and produce C and D.

We will take the incoming and outgoing particle wavefunctions as plane waves, so that in the momentum space we get particles with definite momenta. This description is fine if the underlying theory is a free-field theory, *i.e.*, it does not contain any interaction, but in that case it will be a very boring theory as well: nothing happens. We will need interactions among various particles. In this case, a plane wave description holds only at far away from the interaction region, also known as the asymptotic limit; it should not hold where the interaction takes place. But we have nothing more sophisticated, so we have to do with this. The price we must pay is that in most of the cases, and certainly in all the cases that we will discuss, there is no way to get the exact solution. But this is a familiar situation in quantum mechanics as well, and tools like perturbation theory were developed precisely for this. Well, we will also do a perturbation theory, and truncate the solution somewhere when we either reach the required accuracy or when our patience runs out. Unfortunately, not all field theories can be dealt perturbatively; a notorious example is the theory of strong interactions, quantum chromodynamics, in the low-energy limit.

We will perform the entire calculation in the momentum space. This simplifies the treatment tremendously, and must be considered the best legacy of Richard Feynman to science. Let us see what may go into the expression.

First, the probability of scattering, given by the amplitude of the process squared. Later we will calculate this for several important processes. We denote this by $|\mathcal{M}|^2$, taking the modulus because we are interested only in the magnitude and not the associated phase. Physics lives only in $|\mathcal{M}|^2$, but we need some more bookkeeping. First, we have to multiply this by the available phase space. The larger the available phase space is (for the particles C and D), the larger the cross-section. The phase space, dQ, is written as

$$dQ = \frac{d^3p_C}{(2\pi)^3 2E_C}\frac{d^3p_D}{(2\pi)^3 2E_D}(2\pi)^4\delta^{(4)}(p_A+p_B-p_C-p_D)\,. \tag{140}$$

Let us try to understand the various factors. We use $\hbar = 1$, so the volume of a cell in the 3-dimensional phase space is just $(2\pi)^3$, and we normalise to $2E$ particles per unit volume, so the first fraction gives the number of phase cells in the interval $\mathbf{p}_C$ to $\mathbf{p}_C + \mathbf{dp}_C$ for one C-type particle; similarly for the second fraction.

For the plane waves associated with the particles, the position dependence resides only in the exponential factors: $\exp(-ip_A x)$ and $\exp(-ip_B x)$ for A and B, and $\exp(ip_C x)$ and $\exp(ip_D x)$ for C and D. The signs are opposite, because C and D are outgoing while A and B are incoming, and we assume the interaction to take place at some coordinate x. But x is arbitrary and we must integrate over x; this gives the $(2\pi)^4$ times the four-dimensional delta function. Physically, this is nothing but the energy-momentum conservation requirement.

But there is a problem with $|\mathcal{M}|^2 dQ$; it is not something intrinsic. If the number of particles in the initial beam goes up, so will the number of particles in the final state. So we need to normalise it with the *initial flux* F. Suppose we are observing the collision for a fixed-target experiment where the particles A moving with a velocity $\mathbf{v}_A$ strike the static B particles. Number of A particles crossing unit area per unit time is $|\mathbf{v}_A| 2E_A$ and number of B particles per unit area is $2E_B$, so to get the cross-section, *i.e.*, the probability that one A hits one B per unit time, we have to divide by F, which in the laboratory frame is

$$F = |\mathbf{v}_A| 4E_A E_B \,. \tag{141}$$

While it is not exactly straightforward to derive an expression for F in the centre-of-mass frame (see Peskin and Schroeder for the derivation), let us just quote the result:

$$F = 4p_i\sqrt{s}\,, \tag{142}$$

where $p_i = |\mathbf{p}_A| = |\mathbf{p}_B|$. We get the total cross-section σ by integrating

$$d\sigma = \frac{1}{F}|\mathcal{M}|^2 dQ \tag{143}$$

over the variables. One can also obtain the distribution with respect to any variable by integrating out the rest. Remember: *all quantities that you do not observe must be integrated upon.* Suppose you are observing the collision in the centre-of-mass frame; then the final-state 3-momentum $p_f = |\mathbf{p}_C| = |\mathbf{p}_D|$ is fixed and p_f is no longer a free variable. If you do not observe D, integrate over d^3p_D, and what you get is a differential distribution of C with respect to its angular variables θ_C and ϕ_C; this gives $d\sigma/d\Omega$.

Let us work out an important case. In the centre-of-mass frame, after performing the d^3p_D integration in the expression of (140), we have

$$dQ = \frac{1}{4\pi^2}\frac{d^3p_C}{4E_C E_D}\delta(E_A + E_B - E_C - E_D)\,, \tag{144}$$

where the integration also takes out the 3-momentum δ-function. We can write $d^3p_C = p_f^2 dp_f d\Omega$. Also, $\sqrt{s} = E_A + E_B = E_C + E_D = \sqrt{m_C^2 + p_f^2} + \sqrt{m_D^2 + p_f^2}$. Therefore

$$\frac{d\sqrt{s}}{dp_f} = p_f\left(\frac{1}{E_C} + \frac{1}{E_D}\right)\,, \tag{145}$$

and

$$dQ = \frac{1}{16\pi^2}p_f\left(\frac{1}{E_C + E_D}\right) d\Omega\, d\sqrt{s}\, \delta(\sqrt{s} - E_C - E_D) = \frac{1}{4\pi^2}\frac{p_f}{4\sqrt{s}}\, d\Omega\,. \tag{146}$$

Putting the expression for F, we obtain

$$d\sigma = \frac{1}{64\pi^2 s}\frac{p_f}{p_i}|\mathcal{M}|^2\,. \tag{147}$$

The cross-section σ is an intrinsic quantity (but not Lorentz-invariant). To get the number of events with C and D in the final state, we have to multiply it with some quantity having the dimension of (area)$^{-2}$. This is known as the *integrated luminosity* and gives the number of collisions per unit area. The more basic unit, of course, is the *luminosity*, giving the number of collisions per unit area per unit time, which, when multiplied by the cross-section, gives the rate of production of the final state particles.

Suppose a collider has an integrated luminosity of 10 fb^{-1}, which means 10 collisions per femtobarn, or 10^{16} collisions per barn (1 barn $= 10^{-24}$ cm^2). If the cross-section of an event is 1 pb $= 10^{-12}$ barn, there will be 10^4 such events. If the luminosity is attained in one year of running, which for operational purposes can be taken as 10^7 seconds (the collider needs some rest!), there will be one event per 1000 seconds of running, or one in every 20 minutes.

One, however, cannot observe all these events. Some of them will be missed by the detector; some will be wrongly identified as well. The detection efficiency varies from channel to channel. If the detection efficiency is 5%, one expects only 500 events per year.

Q. Check that $d^3\mathbf{p}/E$ is indeed Lorentz-invariant.

4.5.4 Decay width

For the decay $A \to B + C + \cdots$, one can have a similar master formula as we have found for scattering. The amplitude squared $|\mathcal{M}|^2$ is to be multiplied by the phase space factor dQ, which is a product of individual phase spaces for all n particles in the final state:

$$dQ = (2\pi)^4 \delta^{(4)}(p_A - \sum_{i=1}^{n} p_i) \Pi_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \,. \tag{148}$$

Similarly, the initial flux is just replaced by the number of particles A per unit volume, *i.e.*, $2E_A$. If the particles decays at rest, $E_A = m_A$.

Consider the decay $A \to B + C$ where A was at rest. From (146), using $\sqrt{s} = m_A$, we obtain

$$d\Gamma = \frac{p_f}{32\pi^2 m_A^2}\, |\mathcal{M}|^2 \, d\Omega \,. \tag{149}$$

The total decay width Γ is the sum of the decay widths in all channels. The lifetime of the particle is

$$\tau = \frac{1}{\Gamma} \,. \tag{150}$$

The *partial width* or *branching ratio* (BR) in any particular channel i is given by

$$BR(A \to i\text{th channel}) = \frac{\Gamma_i}{\Gamma} \,. \tag{151}$$

Q. Suppose the particle A can decay into three channels. For channel 1, the BR is 0.60 ± 0.12 and for channel 2, the BR is 0.35 ± 0.10. If all the measurements are uncorrelated (*i.e.*, a maximum in channel 1 does not necessarily mean a minimum in channel 2), what is the expected BR in channel 3?

4.5.5 Collider experiments

Most of the high-energy experiments (and even low-energy nuclear experiments too) are of collider type; two particle collide, some new things are created, and we study their properties. Or, in some cases, two particles collide, and we hope that something new will turn up, and we search for that.

Most of the colliders are either electron-positron or proton-proton (or proton-antiproton). There have been electron-proton collider experiments, and there are proposals even for photon-photon collisions, but we will not discuss them here. The *ep* collisions are done to study the internal structure of the proton, about which you will learn later.

The collider experiments can either be of a fixed-target type, where the beam of particle A collides on a fixed target of B, or of moving-target type, where the beams of A and B come from opposite direction and collide head-on. For most of the cases, the latter type of collisions are done in such a way as to keep the CM of the system at rest (thus, if A and B are identical particles, or antiparticle of one another, the beams would have the same energy). However, this need not be the case always.

Consider the fixed-target experiments first. We can write the four-momenta of A and B as

$$p_A = (E, 0, 0, p)\,, \quad p_B = (m_B, 0, 0, 0)\,, \tag{152}$$

where $E = \sqrt{p^2 + m_A^2}$. The maximum energy available for the creation of new particles is $\sqrt{s} = [(E + m_B)^2 - p^2]^{1/2}$. But $E^2 - p^2 = m_A^2$, so if the particle A is so energetic that $E \gg m_A, m_B$, we can write

$$\sqrt{s} = \sqrt{2 m_B E}\,. \tag{153}$$

Next consider a moving-target experiment keeping the CM at rest. We again write

$$p_A = (E_A, 0, 0, p)\,, \quad p_B = (E_B, 0, 0, -p)\,. \tag{154}$$

If both the particles are relativistic, $E_A \approx E_B = E$, and

$$\sqrt{s} = 2E\,. \tag{155}$$

This shows that the energy reach for moving-target experiments can be much higher than fixed target experiments. The advantage of the fixed-target experiments is that they are easier to design; it is always easier to shoot a fixed target than a flying bird. If the energy needed is not high (like in the nuclear physics experiments where the main goal is to study the excited states of the target or daughter nuclei), one goes for a fixed-target experiment, but if it is high $\sqrt{s}$ that is needed, one has to go for moving-target experiments.

Q. The last run of the Large Electron-Positron collider (LEP-2) was a moving-target fixed-CM collision with $\sqrt{s} = 210$ GeV. Estimate $\sqrt{s}$ if a positron beam of identical energy made a head-on collision with fixed-target electrons. Given, $m_e = 5.1 \times 10^{-4}$ GeV.
Q. The Large Hadron Collider (LHC) is a pp collider, running at present with both proton beams at 4000 GeV. What would have been the necessary energy of a proton beam to create the same $\sqrt{s}$ in a fixed-target pp collision?
Q. Consider an asymmetric e^+e^- collider where the beam energies are different. First, convince yourself that the CM is not at rest in the laboratory frame. Suppose both the beams are relativistic, so that $p_A = (E_A, 0, 0, E_A)$ and $p_B = (E_B, 0, 0, E_B)$. Show that $\sqrt{s} = 2\sqrt{E_A E_B}$.
Q. The asymmetric KEK-B collider hit 3.5 GeV positrons with 8 GeV electrons. This produces, as a resonance, the $\Upsilon(4S)$ meson, which is a $b\bar{b}$ bound state. If the mass of $\Upsilon(4S)$ is 10.5794 GeV, what is its velocity in the laboratory frame? [*Hint: Use* $E = m\gamma$.]

The next question is, when to use an e^+e^- collider and when a pp (or $p\bar{p}$) collider? The question can be rephrased: when do we use a narrow focussed intense beam of light and when a broad searchlight? The answer is easy: when we need to search something but not sure where it can be, we use a searchlight; when we know the object we are looking for and know where it lies, but want to get more detailed information about it, we use a narrow beam.

The hadron colliders are like these broad searchlights. The reason is that the proton is not an elementary particle; it consists of three quarks in the simplest picture, but actually it is much more complicated. There are, of course, these three quarks known as valence quarks, but there are a lot of quark-antiquark pairs (known as sea quarks), and a lot of gluons. A general name for all these constituents is *partons*. When two protons collide with each other, they get smashed, and the partons interact. That means a lot of partonic subprocesses, like qq, $q\bar{q}$, qg, gg, that may lead to final-state particles.

Unfortunately we do not know how much energy an individual parton carries. We know the density of any specific type of partons as a function of energy, and to get the cross-section, one has to convolute the parton level cross-section with these parton distribution functions. For example, the differential cross-section for $q\bar{q} \to q'\bar{q'}$ (where q and q' are different quarks) is given by

$$\frac{d\sigma}{d\Omega}(q\bar{q} \to q'\bar{q'}) = \frac{\alpha_s^2}{9s}\frac{t^2+u^2}{s^2}\,, \tag{156}$$

where $\sqrt{4\pi\alpha_s}$ is the gauge coupling g_3 for the strong interaction. The Mandelstam part is identical to what one gets for $e^+e^- \to \mu^+\mu^-$ in QED, but the overall factor is different because of the properties of the SU(3) group. However, for hadronic collisions, we use $\hat{s}$, $\hat{t}$, and $\hat{u}$ to denote parton-level Mandelstam variables. Thus, if the incoming partons carry x_1 and x_2 fractions of the momentum of the proton, the CM energy $\hat{s} = x_1x_2s$. We write

$$\sigma_{had} = \int dx_1\ dx_2\ f_i(x_1)f_j(x_2)\hat{\sigma}\,, \tag{157}$$

where $\hat{\sigma}$ is the partonic level cross-section. The probability that the i-th incoming parton will have a momentum fraction between x_1 and $x_1 + dx_1$ is given by the *parton distribution function* $f_i(x_1)$, and similarly for $f_j(x_2)$.

In short, any hadronic collision leads to a huge lot of final-state particles, most of them being junk (i.e., those we don't want to investigate). To search something new, we have to scan through this junkheap, over billions and billions of such collisions. But we have done this before, and are doing so again at the LHC.

The situation is simpler in a collision with elementary point-like particles in the initial state. The energy is given, and we know $\sqrt{s}$, so we fix $\sqrt{s}$ to the point we want. For example, if we want to produce the weak neutral Z boson, which can be produced as a resonance [10], we may tune the beam energies to $\sqrt{s}/2$ each, and the machine should produce real Z bosons copiously, acting as a Z-factory. This is precisely what happened at the first round of LEP. If we want to produce a pair of particles, say W^+W^-, we tune each beam energy to half the energy needed.

Thus, the e^+e^- collider works only if we know what we are looking for. There is another limitation; it is hard to build a circular e^+e^- collider with an energy significantly more than what the last run of LEP achieved, i.e., about 210 GeV. The reason is synchrotron radiation; it goes as $1/m^4$, and as the energy increases, the loss due to synchrotron radiation becomes more than the energy pumped into the beams per revolution, so it becomes a loss-making concept. The way out is to build a linear collider, but for that we need a really long machine.

---

[10] This is called a resonance because the cross-section shows a sharp bump. It comes from the fact that the amplitude $e^+e^- \to f\bar{f}$, where f is any fermion, when mediated by the Z boson, has the propagator of the form $1/(s - m_Z^2 + im_Z\Gamma_Z)$. Γ_Z is the decay width of the Z; this is a suitable form to write the propagator for an unstable particle, known as the *Breit-Wigner form*, but this is nothing but the $+i\varepsilon$ prescription for the propagator. How do we get the Breit-Wigner form? For a stable particle, one can write the evolution as $\psi(t) = \exp(-imt)\psi(0)$, so that $|\psi(t)|^2 = |\psi(0)|^2$. For an unstable particle, the practice is to replace m by $m - i\Gamma/2$, so that $|\psi(t)|^2 = \exp(-\Gamma t)|\psi(0)|^2$, the exponential decay. Replace m by $m - i\Gamma/2$ in the propagator, and use the approximation $\Gamma \ll m$, to get the Breit-Wigner form. At $s = m_Z^2$, the first two terms sum to zero, and the last term controls the amplitude. This being small, there is a sharp rise in the cross-section. This is completely analogous to anomalous scattering of electromagnetic waves.

Now, this is a very simplistic picture of an e^+e^- machine. There are two effects that cause a spread in the final-state total energy. The first is known as *initial state radiation* (ISR), where the incoming electron or positron radiates an energetic photon in the same direction of its motion (this is called a hard collinear photon), and thus the energy is degraded to $\sqrt{\tilde{s}} = \sqrt{s} - E_\gamma$. This can be pretty well computed in QED. The second effect is important only if the particle densities in the beams are high. In that case, the electron radiates a photon just before a collision under the influence of the electromagnetic field of the positron (or vice versa). This is called *beamstrahlung*, and this is also quite accurately computable [11].

4.6 π^- decay

Consider the decay

$$\pi^-(q) \to \mu^-(p) + \bar{\nu}_\mu(k) \tag{158}$$

whose amplitude must be of the form

$$i\mathcal{M} = i\frac{G_F}{\sqrt{2}}\bar{u}(p)\gamma^\mu(1-\gamma_5)v(k) \times \text{hadronic contribution}\,. \tag{159}$$

The hadronic contribution is unknown because d and $\bar{u}$ are not free; the confinement dynamics is not known. But this must be a Lorentz vector, as the current carries the Lorentz index:

$$\text{hadronic contribution} = \langle 0|\bar{u}\gamma_\mu(1-\gamma_5)d|\pi^-\rangle = if_\pi q_\mu\,. \tag{160}$$

Two points have to be noted. First, we have assumed that the pion goes to a vacuum and out of that vacuum come the leptons. Technically we should take all intermediate states with quantum numbers of the vacuum, but this is a pretty good appproximation as at this energy, no other state contributes. Second, the final result must be proportional to q_μ, because that is the only possible four-vector that we have (if this were a vector meson, we could have used the polarization too). The proportionality constant cannot be calculated from first principles, so we have dumped that ignorance into the *decay constant* f_π, which must be obtained experimentally.

This gives

$$\begin{aligned}\mathcal{M} &= i\frac{G_F}{\sqrt{2}}f_\pi q_\mu\left[\bar{u}(p)\gamma^\mu(1-\gamma_5)v(k)\right]\\ &= i\frac{G_F}{\sqrt{2}}f_\pi\left[\bar{u}(p)(\not{p}+\not{k})(1-\gamma_5)v(k)\right]\\ &= i\frac{G_F}{\sqrt{2}}f_\pi m_\mu\bar{u}(p)(1-\gamma_5)v(k)\,,\end{aligned} \tag{161}$$

where we have used the Dirac equation $\not{k}v(k) = \bar{u}(p)(\not{p}-m_\mu) = 0$.

It is left as an exercise to show that

$$|\mathcal{M}|^2 = 4G_F^2 f_\pi^2 m_\mu^2 (p.k)\,. \tag{162}$$

---

[11] Something interesting may happen due to ISR and beamstrahlung. Suppose the initial beams have $\sqrt{s} > m_Z$. You do not expect a resonance. But there can be photons radiated off, and the degraded energy may fall on the resonance: $\sqrt{\tilde{s}} = m_Z$. While the amplitude for such a process will be down by a factor α, the resonance cross-section is so large that you will see a significant number of events at the resonance; there will be the Z peak in the cross-section plot. This is known as a *radiative return.*

But $q = p + k$, so

$$p.k = \frac{1}{2}(q^2 - p^2 - k^2) = \frac{1}{2}(m_\pi^2 - m_\mu^2)\,. \tag{163}$$

Using (149), one gets the differential decay width $d\Gamma$, but there is no angular dependence, so the $d\Omega$ integration just gives 4π. The final state momentum is given by (133):

$$|\mathbf{p}| = (m_\pi^2 - m_\mu^2)/2m_\mu\,. \tag{164}$$

Putting everything together, we get the decay width as

$$\Gamma = \frac{G_F^2}{8\pi} f_\pi^2 m_\pi m_\mu^2 \left(1 - \frac{m_\mu^2}{m_\pi^2}\right)^2\,. \tag{165}$$

Note the factor of m_μ^2. This shows that the decay to $\mu\bar{\nu}_\mu$ is much more probable than $e\bar{\nu}_e$, even though the phase space for the latter decay is larger. The physics is interesting.

Pion is a spin-0 object. So the total angular momentum of the lepton pair must be zero. This means that while the leptons come out back-to-back (in the rest frame of the pion), the spins are also pointing opposite. The $\bar{\nu}$ is a right-chiral field, so the spin and the momentum must be parallel. The same will be true for the muon, so it must have right helicity. At the same time, it must be left-chiral, because only the left-chiral particle fields can have weak interaction. The more massive a fermion is, the more chance it has that the chirality will not be the same as the helicity. In other words, the muon comes out left-chiral, but there must be a mass term inserted somewhere.

Q. Suppose the object that couples to the leptonic current is not a spin-1 object (the virtual W), but a spin-0 object whose coupling is proportional to the μ mass. Show that you would have got the same result.

4.7 Muon decay

The decay $\mu \to e + \bar{\nu}_e + \nu_\mu$ can be treated with the Fermi theory, but the calculation is more complicated as we have three particles in the final state. We just quote the final result:

$$\Gamma = \frac{G_F^2 m_\mu^5}{192\pi^3} \tag{166}$$

if we neglect the final state lepton masses.

One can include the corrections coming from $m_e \neq 0$. This being the only decay channel of muon, the decay gives a very precise measurement of G_F.

Q. Given, $G_F = 1.16637 \times 10^{-5}$ GeV^{-2} and $m_\mu = 105$ MeV. What is the lifetime of muon in s?

5 The Electroweak Theory

Let us reiterate our main problem of weak interaction: how to have massive gauge bosons without breaking the gauge invariance? One can write a term $-m^2 A_\mu A^\mu$, but it explicitly spoils gauge invariance; under the gauge transformation $A_\mu \to A_\mu + \partial_\mu \lambda$, the mass term picks up extra pieces. For QED we did not face the problem, as photon was massless, but we must tackle this issue if we wish to form a gauge theory of weak interactions.

5.1 Spontaneous breaking of a discrete symmetry

Consider a real scalar field $\phi(x)$ with a Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - V(\phi)\,, \tag{167}$$

where

$$V(\phi) = \frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda\phi^4\,. \tag{168}$$

This is something that we know; the first term of $V(\phi)$ is the mass term of the field, and the second term is the standard four-point interaction. Of course, we expect $m^2 > 0$, and we should also have $\lambda > 0$ so that the potential does not become unbounded from below as we go to large-ϕ limit. Now recall the main result of interacting field theory: we must do a perturbation calculation. The minimum energy state, the vacuum, is obtained by minimising the potential, and all we can do is to get a perturbation expansion around this minimum. The necessary requirement for this is $|\lambda| < 1$.

Minimising $V(\phi)$, with $\partial V/\partial\phi = 0$ and $\partial^2 V/\partial\phi^2 > 0$, we get only one solution: $\phi = 0$. The minimum is at $\phi = 0$, so that is the vacuum configuration, and the *vacuum expectation value* (VEV) $\langle\phi\rangle = \langle 0|\phi|0\rangle = 0$, as the field operator ϕ can be expressed as a linear combination of the creation and the annihilation operators.

Now suppose $\lambda > 0$ but $m^2 < 0$. Then m^2 is just some parameter of the theory and cannot be directly related with the mass of the field quantum. The Lagrangian is symmetric under $\phi \to -\phi$, which is a discrete symmetry, called $\mathbf{Z}_2$.

The minimisation condition $\partial V/\partial\phi = 0$ now yields

$$\phi(m^2 + \lambda\phi^2) = 0\,. \tag{169}$$

But the second derivative at $\phi = 0$ is m^2, which is negative, so $\phi = 0$ is a local maximum! Thus it is an unstable equilibrium point and no perturbation expansion can be done around $\phi = 0$. However, there are two other minima, given by

$$\phi = \pm\sqrt{-\frac{m^2}{\lambda}}\,. \tag{170}$$

The field configuration must live in any one of these two minima. We have no a priori way to say which configuration it will choose, but it will choose any one of these; both are equally probable.

Suppose a large number of people are sitting at a round table for dinner, and glasses are placed exactly at the midpoints of two persons. This is a perfectly symmetric configuration for the glasses, but the person who first chooses to pick up her glass breaks the symmetry. Either she chooses the left-hand side or the right-hand side glass, and that fixes the choice for all the people (otherwise, some will get two glasses, and some won't get even one). There was a symmetry, but once that is broken, the entire configuration goes into that particular symmetry-broken phase.

This is known as a *spontaneous symmetry breaking* (SSB). The original Lagrangian had a $\mathbf{Z}_2$ symmetry, but the configuration must go in either of the two minima. If we call

$$v = \sqrt{-\frac{m^2}{\lambda}}\,, \tag{171}$$

then there are two vacua which are equally probable: one with $\langle\phi\rangle = v$ and another with $\langle\phi\rangle = -v$. Once a particular vacuum is chosen, the $\mathbf{Z}_2$ symmetry is spontaneously broken. This is then the definition of SSB: The Lagrangian is symmetric under some symmetry but the vacuum is not [12].

Let us assume that the system has chosen $\langle\phi\rangle = v$. We must perform perturbation around this point only. So we will define another field h:

$$h = \phi - v \quad \Rightarrow \phi = h + v\,, \tag{172}$$

so that the VEV of h, $\langle h\rangle$, is zero. In this vacuum, the potential reads (the kinetic term does not change)

$$\begin{aligned} V(h) &= \frac{1}{2}m^2(h+v)^2 + \frac{1}{4}\lambda(h+v)^4 \\ &= \frac{1}{4}\lambda h^4 + \lambda v h^3 + \left(\frac{1}{2}m^2 + \frac{3}{2}\lambda v^2\right)h^2 + (\lambda v^3 + m^2 v)h + \frac{1}{2}m^2v^2 + \frac{1}{4}\lambda v^4 \\ &= \frac{1}{4}\lambda h^4 + \lambda v h^3 + (-m^2)h^2 - \frac{1}{4}\lambda v^4\,, \end{aligned} \tag{173}$$

where we have used the minimisation condition $m^2 + \lambda v^2 = 0$.

Eq. (173) is interesting. It tells you that in the symmetry-broken phase,

- The four-point interaction is present with the original strength. This is the only term with a dimensionless coupling, so the renormalisation procedure does not depend on whether the symmetry has been broken.
- There is a cubic term $\lambda v h^3$ — this is the signature of SSB, a term that does not respect $\mathbf{Z}_2$.
- There is no term linear in h, the so-called tadpole terms.
- There is a constant negative contribution to the potential; the minimum of the potential has been shifted downwards. But this need not worry us, because we never measure the actual potential, only the energy level differences, and as long as it is bounded from below, we are safe. We also know that any such constant contribution can be eliminated by a normal ordering of the fields.
- The most important one: there is a proper mass term, quadratic in h! As $m^2 < 0$, this comes with a positive sign, and the mass of the field is $\sqrt{2m^2}$.

So what's the lesson? In order to get the physical spectrum, one must expand around the minimum [13]. Our next task is to do the analysis for a continuous symmetry.

---

[12] Note that this is a classical description. A quantum field can always be expressed in terms of creation and annihilation operators, and has zero VEV by definition. So, what we are doing is a quantum description of the field around a classical minimum.

[13] Is it certain that the entire field configuration will go to the same symmetry-broken phase? In our dinner-table analogue, is it a must that once someone picks up the left-hand glass, everyone will do so? The answer is no; it can easily happen that at the same time, A picks up her left-hand glass and B picks up his right-hand one. If you are imaginative enough, you can think the table to be so big, and so full of foodstuff, that B cannot see what A is doing. This will necessarily leave at least one person embarrassed with two glasses, and at least one person angry with no glasses at all. Such boundaries among different vacua are known as domain walls. There was an SSB in the very early stage of the universe; were such domain walls formed? As far as we know, nothing like that happened. Why? What told the different parts of the baby universe, which are causally disconnected, to follow the same SSB pattern? The very interesting answer to this question is unfortunately outside the scope of this book, but you can look at any standard cosmology textbook, or search the web for inflationary universe.

Q. What would be the potential if we choose the other vacuum $\langle\phi\rangle = -v$? Show that the cubic term will flip its sign.

5.2 Spontaneous breaking of a continuous symmetry: Goldstone theorem

Now let us consider the Lagrangian for two real scalar fields ϕ_1 and ϕ_2:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi_1)(\partial^\mu\phi_1) + \frac{1}{2}(\partial_\mu\phi_2)(\partial^\mu\phi_2) - V(\phi_1,\phi_2)\,, \tag{174}$$

where

$$V(\phi_1,\phi_2) = \frac{1}{2}m^2(\phi_1^2+\phi_2^2) + \frac{1}{4}\lambda(\phi_1^2+\phi_2^2)^2\,. \tag{175}$$

The Lagrangian is symmetric under a two-dimensional rotation or $O(2)$, *i.e.*, any transformation that keeps $\phi_1^2+\phi_2^2$ invariant:

$$\xi_1 = \phi_1\cos\theta + \phi_2\sin\theta\,,\quad \xi_2 = -\phi_1\sin\theta + \phi_2\cos\theta\,. \tag{176}$$

We can also write the Lagrangian in terms of the combinations

$$\phi^+ = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)\,,\quad \phi^- = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2)\,, \tag{177}$$

as

$$\mathcal{L} = (\partial_\mu\phi^+)(\partial^\mu\phi^-) - m^2\phi^+\phi^- - \lambda(\phi^+\phi^-)^2\,. \tag{178}$$

We will again choose $\lambda > 0$ but $m^2 < 0$ to get SSB. The minimisation condition $\partial V/\partial\phi_1 = 0$ gives

$$\left(m^2 + \lambda(\phi_1^2+\phi_2^2)\right)\phi_1 = 0\,. \tag{179}$$

There is a similar equation for ϕ_2. Again, just as before, $\phi_1 = \phi_2 = 0$ is a local maximum and is not suitable for a perturbation expansion. Thus, in the stable configuration, one must have

$$\phi_1^2+\phi_2^2 = -\frac{m^2}{\lambda} = v^2\,. \tag{180}$$

This is a circle in the ϕ_1-ϕ_2 plane, every point of which is a possible vacuum and the system can go into any one of these infinite number of vacua. We have to choose a particular configuration, just as we did in the discrete case, and let us choose

$$\langle\phi_1\rangle = v\,,\quad \langle\phi_2\rangle = 0\,. \tag{181}$$

We will again shift ϕ_1 to $h = \phi_1 - v$ so that $\langle h\rangle = 0$, but there is no need to shift ϕ_2. The kinetic terms are always invariant under these constant shifts, and so let us focus on the potential term:

$$\begin{aligned} V(h,\phi_2) &= \frac{1}{2}m^2\left[(h+v)^2+\phi_2^2\right] + \frac{1}{4}\lambda\left[(h+v)^2+\phi_2^2\right]^2 \\ &= -\frac{1}{4}\lambda v^4 - m^2h^2 + \lambda v h^3 + \lambda v h\phi_2^2 + \frac{1}{2}\lambda h^2\phi_2^2 + \frac{1}{4}\lambda\left(h^4+\phi_2^4\right)\,, \end{aligned} \tag{182}$$

where we have again used the minimisation condition $m^2 + \lambda v^2 = 0$.

Time to take stock of what we get. Apart from those that we found in the discrete symmetry case, here we see:

1. The field h gets a mass, $\sqrt{-2m^2} = \sqrt{2\lambda v^2}$.
2. The field ϕ_2 does not have a mass term. Even after SSB, this remains massless.

It is not difficult to understand why ϕ_2 still remains massless. In the ϕ_1 direction, the minimum is at the bottom of a trough, and an excitation means a climb along the potential well. This requires some energy, and so the vacuum-shifted ϕ_1, which is h, gets a mass. Along the ϕ_2 direction — which is a tangent on the circle of potential minimum — a shift does not involve a climb, and so no energy is needed for the excitation. This keeps ϕ_2 massless.

This is an example of a very important theorem in field theory, known as *Goldstone theorem*: If a continuous symmetry is spontaneously broken, there exists massless scalars in the spectrum, and the number of massless scalars is equal to the number of broken generators (of the concerned symmetry group). Such scalars are called *Goldstone bosons*. The number of broken generators is just the difference between the number of unbroken generators of the original unbroken symmetry and the number of generators of the symmetry that remains after breaking. Here, initially there was one generator of O(2), and after the symmetry is broken, there is no continuous symmetry, so the number is zero, and hence we have one Goldstone boson.

We can go for more complicated scalar structure. Suppose we have a complex scalar doublet Φ:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_0 + i\phi_3 \end{pmatrix} , \tag{183}$$

and the potential is

$$V(\Phi) = m^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2 = \frac{1}{2} m^2 \sum_{i=0}^{3} \phi_i^2 + \frac{1}{4}\lambda \left(\sum_{i=0}^{3} \phi_i^2\right)^2 . \tag{184}$$

If $m^2 < 0$, the minimisation condition is

$$m^2 + \lambda \sum_{i=0}^{3} \phi_i^2 = 0 . \tag{185}$$

Let us take $\langle \phi_0 \rangle = v$ and $\langle \phi_i \rangle = 0$ for $i = 1$ to 3. The vacuum-shifted potential in terms of $h = \phi_0 - v$ and ϕ_i is

$$V(h, \phi_i) = \frac{1}{2} m^2 \left[(h+v)^2 + \sum_{i=1}^{3} \phi_i^2\right] + \frac{1}{4}\lambda \left[(h+v)^2 + \sum_{i=1}^{3} \phi_i^2\right]^2 . \tag{186}$$

It is easy to check that h gets a mass of $\sqrt{-2m^2}$ and there are three massless Goldstone bosons $\phi_{1,2,3}$. Initially, the symmetry was $O(4)$ with six generators, and after breaking, the remaining symmetry is only $O(3)$ with three generators, so there should be three Goldstone bosons in the spectrum, as there are.

Q. Consider the potential of (175). Assign the VEVs as $\langle \phi_1 \rangle = v_1$, $\langle \phi_2 \rangle = v_2$ so that $v_1^2 + v_2^2 = v^2$. Vacuum shift both the fields to h_1, h_2 and find the mass matrix in the h_1-h_2 basis. Diagonalise to get the stationary states. Show that one of them has mass of $\sqrt{2\lambda v^2}$ while the other is massless.

5.3 Spontaneous breaking of an abelian gauge symmetry: Higgs mechanism

Next let us focus on a gauge symmetry. Suppose the gauge group is abelian, with only one gauge boson, and the ordinary derivative is replaced by the covariant derivative:

$$\partial_\mu \to D_\mu = \partial_\mu + igA_\mu . \tag{187}$$

Consider the simplest continuous symmetry with two real scalar fields. The potential is the same as (175), but the kinetic term looks like

$$\mathcal{L}_{kin} = [(\partial_\mu - igA_\mu)\phi_1\ (\partial^\mu + igA^\mu)\phi_1] + [(\partial_\mu - igA_\mu)\phi_2\ (\partial^\mu + igA^\mu)\phi_2] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,. \tag{188}$$

The minimisation condition does not change, and we again shift $\phi_1 \to h + v, \phi_2 \to \phi_2$. The kinetic term looks like

$$\mathcal{L}_{kin} = (D_\mu h)^\dagger(D^\mu h) + (D_\mu\phi_2)^\dagger(D^\mu\phi_2) + 2g^2v^2A_\mu A^\mu h^2 + g^2v^2A_\mu A^\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,. \tag{189}$$

Something strange has happened. We have not broken the gauge symmetry, but the gauge field has acquired a mass term $g^2v^2A_\mu A^\mu$, *i.e.*, the field now has a mass gv.

But that is clearly unphysical, as a massive gauge boson has three degrees of freedom, compared to two for a massless one, and shifting the vacuum cannot generate an extra degree of freedom. Thus, one of the degrees of freedom is spurious, and we should cleverly choose a gauge where the spurious degree can be identified. To this end, write the potential as in (178), in terms of ϕ^+ and ϕ^-. Note that the gauge transformation is equivalent to

$$\phi^+(x) \to \exp[i\alpha(x)]\phi^+(x)\,, \quad \phi^-(x) \to \exp[-i\alpha(x)]\phi^-(x)\,. \tag{190}$$

In the vacuum-shifted configuration, we can write

$$\phi^+ = \frac{1}{\sqrt{2}}\,[h + v + i\phi_2] = \frac{1}{\sqrt{2}}\,[h + v]\exp(i\phi_2/v)\,, \tag{191}$$

keeping only the terms linear in the fields. We have a freedom to choose $\alpha(x)$; we can choose it in such a way that it exactly cancels the exponential part: $\alpha(x) = -\phi_2(x)/v$. Thus, in the gauge-transformed potential, there is no ϕ_2, and we can remove this degree of freedom from our analysis. This particular choice of gauge, where only the physical particles remain in the spectrum, is called the *unitary gauge*. This saves the situation: before SSB, there were four degrees of freedom, two for the gauge boson, and two for the two real scalars, and after SSB, there are again four, three for the massive gauge boson and one for h. We say that the gauge boson has become massive by eating up the Goldstone boson, and we expect the longitudinal degree of freedom of A_μ to reflect some properties of the vanished Goldstone. This is known as the *Higgs mechanism*, although other people, like Anderson, Englert and Brout, and Guralnik, Hagen, and Kibble discussed the mechanism at about the same time as Peter Higgs did [14].

However, when computing the amplitude for a process, one generally keeps to some unphysical gauge, where the gauge boson is massive and the Goldstone boson is also present. This is just for the reason of simplicity: the gauge propagator looks simpler here than in the unitary gauge. For example, in the unphysical 't Hooft-Feynman gauge, the gauge propagator is $-i\eta_{\mu\nu}/k^2 - M^2$, where M is the mass of the gauge boson. In another unphysical gauge, the Landau gauge, the propagator is $-i(\eta_{\mu\nu} - k_\mu k_\nu/k^2)/(k^2 - M^2)$.

How to know whether a gauge is unphysical? Look if the Goldstone is there; in the unitary gauge, the propagator of the Goldstone is zero, and it simply vanishes from the spectrum.

---

[14] Why do we call the celebrated boson after Higgs only? Only Higgs, in his PRL paper, stated the possibility that such a mechanism will result in a spin-0 physical object in the spectrum. This is not even a two-page paper, and only one line about the Higgs boson: "... an essential feature of the type of theory .... is the prediction of incomplete multiplets of scalar ... bosons." But this one line at the last paragraph may well bring a Nobel prize.

5.4 Spontaneous breaking of a non-abelian gauge symmetry

The last task is to extend the Higgs mechanism for a non-abelian gauge group. This is necessary because we have to give masses to three weak gauge bosons, which belong to the adjoint representation of the simplest non-abelian group $SU(2)$. We need three Goldstone bosons, so we start from our old friend, the complex scalar Φ (183), and the potential (184). The covariant derivative is

$$D_\mu = \partial_\mu + i\frac{g}{2}\vec{\tau}.\mathbf{W}_\mu\,, \tag{192}$$

where g is the $SU(2)$ gauge coupling and τs are the Pauli matrices. The kinetic term can be written as

$$(D_\mu\Phi)^\dagger(D^\mu\Phi) = \frac{1}{2}\left[\left(\partial_\mu - i\frac{g}{2}\vec{\tau}.\mathbf{W}_\mu\right)\begin{pmatrix}\phi_1 + i\phi_2\\ h + v + i\phi_3\end{pmatrix}\right]^\dagger \left[\left(\partial^\mu + i\frac{g}{2}\vec{\tau}.\mathbf{W}^\mu\right)\begin{pmatrix}\phi_1 + i\phi_2\\ h + v + i\phi_3\end{pmatrix}\right]. \tag{193}$$

One can separate out the interaction terms:

$$\frac{1}{8}g^2\begin{pmatrix}\phi_1 - i\phi_2 & h + v - i\phi_3\end{pmatrix}\begin{pmatrix}W_{\mu 3} & \sqrt{2}W_\mu^+\\ \sqrt{2}W_\mu^- & -W_{\mu 3}\end{pmatrix}\begin{pmatrix}W_3^\mu & \sqrt{2}W^{\mu +}\\ \sqrt{2}W_1^{\mu -} & -W_3^\mu\end{pmatrix}\begin{pmatrix}\phi_1 + i\phi_2\\ h + v + i\phi_3\end{pmatrix}, \tag{194}$$

where $\sqrt{2}W_\mu^\pm = W_{\mu 1} \mp iW_{\mu 2}$. The gauge boson mass terms can be obtained by putting all scalar fields equal to zero and evaluating (194):

$$\begin{aligned} & \frac{1}{8}g^2\begin{pmatrix}0 & v\end{pmatrix}\begin{pmatrix}W_{\mu 3} & W_{\mu 1} - iW_{\mu 2}\\ W_{\mu 1} + iW_{\mu 2} & -W_{\mu 3}\end{pmatrix}\begin{pmatrix}W_3^\mu & W_1^\mu - iW_2^\mu\\ W_1^\mu + iW_2^\mu & -W_3^\mu\end{pmatrix}\begin{pmatrix}0\\ v\end{pmatrix}\\ = & \frac{1}{8}g^2v^2\left(2W_\mu^+W^{\mu -} + W_{\mu 3}W_3^\mu\right). \end{aligned} \tag{195}$$

From (195), it appears that all the weak gauge bosons — two charged and one neutral — are massive and degenerate:

$$M_W = \frac{1}{2}gv\,. \tag{196}$$

This is a major step forward. The only problem is that the neutral gauge boson Z is a bit more massive than $W^\pm$, so this theory requires some minor modification. In the next chapter, we will see how that modification unifies electromagnetism with weak interaction.

The Higgs mechanism involves two distinct phases: one where the symmetry is unbroken and $v = 0$, and one where the symmetry is broken and $v \neq 0$. From the measured values of the $SU(2)$ gauge coupling g and the mass of W, it appears that $v \approx 246$ GeV. Thus, we can take v as the order parameter between these two phases. We have reasons to believe that in the very early universe, a phase transition occurred from the unbroken phase to the broken phase. It is not sure whether the transition was first order — one where v changed discontinuously across the two phases — or second order, where the transition was continuous.

Q. From (194), find the vertex factors for W^+W^-hh and W^+W^-h vertices

5.5 The electroweak unification

The 'unification' of the electromagnetic and weak forces is the central theme of this subsection. We deal here with the more formal aspects of unification, while the phenomenology will be taken up in the next two chapters. We use the quote in the first sentence deliberately: this is not a true unification in the sense that there is not a single gauge group at some energy which breaks down to the weak gauge group of SU(2) and the electromagnetic gauge group U(1). There may be; in fact there may be a common one which also incorporates the colour SU(3); but we have not yet found any signature of that.

The model of electroweak unification was proposed by Glashow, Salam, and Weinberg. That is why this is also known as the Glashow-Salam-Weinberg (GSW) theory; sometimes the order of S and W is reversed, and sometimes this is loosely called the 'Standard Model', though the SM includes the electroweak gauge groups as well as the unbroken SU(3) of colour.

The GSW model is based on the gauge group SU(2) × U(1). Thus, all states are to be labelled by two quantum numbers, one for the SU(2) and other for the U(1). The SU(2) is known to us; this is the weak interaction gauge group, with three generators that correspond to three gauge bosons, two charged and one neutral. All of them are massive, but we know how to make them massive in a gauge-invariant way. There must exist in the model a complex scalar doublet which gives rise to three Goldstone bosons, to be eaten up by the three weak gauge bosons. The weak isospin projection T_3 is a good quantum number to label the SU(2) states. From Wu's experiment, we know that weak interaction violates parity, and from Goldhaber's experiment, we know that neutrinos are completely left-handed. Neutrinos have only weak interaction — it is a neutral lepton — so the property of neutrinos is a good reflection to the property of the weak interaction. The GSW model assumes that the left-handed fermions $\psi_L = P_L\psi$ are doublets under SU(2), *i.e.*, they live in the fundamental representation. The members of the doublet are connected by charge changing weak interaction; so one doublet is $\begin{pmatrix} \nu_e \\ e \end{pmatrix}$, and another doublet is $\begin{pmatrix} u \\ d \end{pmatrix}$, and similarly for the higher generation leptons and quarks. The right-handed fermions $\psi_R = P_R\psi$ are all singlets, *i.e.*, they do not transform under SU(2). In other words, they do not feel the weak force. That is why we call the gauge group $\mathrm{SU}(2)_L$.

The U(1) is also reminiscent of electromagnetism. But there is a subtle catch. The U(1) generator must commute with all the SU(2) generators, and hence with the raising and lowering operators of the SU(2). That is equivalent to say that both the members of the doublet must have the same U(1) charge. But neutrino and electron do not have the same electric charge; so this U(1) cannot be that of electromagnetism. On the other hand, note that $Q - T_3$ is the same for both the components, and therefore the U(1) charge must be proportional to $Q - T_3$. We take, by convention, the proportionality constant to be 2, and define

$$Y = 2(Q - T_3) \quad \Rightarrow Q = T_3 + \frac{1}{2}Y\,. \tag{197}$$

The *weak hypercharge* Y is the charge of this U(1) and we call this the hypercharge U(1) or simply $\mathrm{U}(1)_Y$. Thus, the GSW gauge group is $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. These two groups are disjoint, with two different coupling constants, g and g' respectively. For example, one can write the covariant derivative as

$$D_\mu = \partial_\mu + i\frac{g}{2}\vec{\tau}.\mathbf{W}_\mu + i\frac{g'}{2}YB_\mu\,. \tag{198}$$

There are four gauge bosons in the theory, all massless to start with. The kinetic terms of the gauge fields are $-\frac{1}{4}\mathbf{W}_{\mu\nu}\mathbf{W}^{\mu\nu}$ and $-\frac{1}{4}B_{\mu\nu}B^{\mu\nu}$ respectively, where

$$\begin{aligned} W^a_{\mu\nu} &= \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\epsilon_{abc}W^b_\mu W^c_\nu\,, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu\,. \end{aligned} \tag{199}$$

Here the index a runs from 1 to 3. Because of the non-abelian nature of SU(2), there is an extra term in $W^a_{\mu\nu}$, which gives rise to 3- and 4-point gauge vertices. The same term is also present in $\mathrm{SU}(3)_c$, where ϵ_{abc} is replaced by the SU(3) structure constants f_{abc}.

The Higgs mechanism gives mass to three of the gauge bosons and one will remain massless. What we will show is

- The neutral gauge boson Z is more massive than the charged ones $W^\pm$, which it should be.
- The massless gauge boson is neutral, and couples with equal strength to the L- and R-fields. Thus, it can be identified with the photon.

A massless gauge boson corresponds to an unbroken gauge symmetry — thus, there is an unbroken U(1), which we can now identify with the U(1) of electromagnetism. Thus the breaking pattern can be written as

$$\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y \to \mathrm{U}(1)_{em}\,. \tag{200}$$

5.6 The mass of the weak gauge bosons

The GSW model has a complex scalar doublet whose symmetry is broken spontaneously, giving three Goldstone bosons. The discussion is completely identical to that in Section 5.4, and the charged weak bosons have the same mass, *viz.*, $\frac{1}{2}gv$. The abelian gauge boson B_μ must be electrically neutral, and so the neutral gauge boson part should be modified.

For the scalar doublet, $Y = 1$. This can be checked from the Q and T_3 quantum numbers of the neutral member: $Q = 0, T_3 = -\frac{1}{2}$. To get the mass terms, it is enough to write the neutral gauge boson part of $(D_\mu\Phi)^\dagger(D^\mu\Phi)$, shifting the vacuum $\phi_0 \to h + v$, and dropping all the scalar fields, keeping only the VEV. This gives

$$\begin{aligned}
&\begin{pmatrix} 0 & \frac{v}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{g}{2}W_{\mu 3} + \frac{g'}{2}B_\mu & 0 \\ 0 & -\frac{g}{2}W_{\mu 3} + \frac{g'}{2}B_\mu \end{pmatrix} \begin{pmatrix} \frac{g}{2}W^\mu_3 + \frac{g'}{2}B^\mu & 0 \\ 0 & -\frac{g}{2}W^\mu_3 + \frac{g'}{2}B^\mu \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \\
= &\ \frac{1}{8}v^2 \left(-gW_{\mu 3} + g'B_\mu\right)\left(-gW^\mu_3 + g'B^\mu\right) \\
= &\ \frac{1}{8}v^2 \begin{pmatrix} W_{\mu 3} & B_\mu \end{pmatrix} \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^\mu_3 \\ B^\mu \end{pmatrix}\,.
\end{aligned} \tag{201}$$

So neither W_3 nor B are stationary states. However, it is obvious that the eigenvalues are $\frac{1}{8}v^2(g^2 + g'^2)$ and zero (the trace and the determinant must remain invariant under a similarity transformation). Suppose we define

$$g = \sqrt{g^2 + g'^2}\cos\theta_W\,, \quad g' = \sqrt{g^2 + g'^2}\sin\theta_W\,, \tag{202}$$

so that the stationary states can be identified as

$$Z_\mu = W_{\mu 3}\cos\theta_W - B_\mu\sin\theta_W\,, \quad A_\mu = W_{\mu 3}\sin\theta_W + B_\mu\cos\theta_W\,. \tag{203}$$

The mass matrix now can simply be written as $\frac{1}{8}(g^2 + g'^2)Z_\mu Z^\mu$,and as Z is a real field, its mass is

$$M_Z = \frac{1}{2}\sqrt{g^2 + g'^2}v\,. \tag{204}$$

In other words,

$$\frac{M_W}{M_Z} = \frac{g}{\sqrt{g^2+g'^2}} = \cos\theta_W\,. \tag{205}$$

This is one of the most important predictions of the GSW model. As we will soon see, the angle θ_W, called the *weak mixing angle* or the *Weinberg angle*, can be very precisely measured, and so can be the masses of W and Z. The relation (205) undergoes some correction due to radiative effects, but nevertheless is satisfied to a remarkable accuracy.

One can reverse (203) and write

$$W_{\mu 3} = Z_\mu \cos\theta_W + A_\mu \sin\theta_W\,, \quad B_\mu = -Z_\mu \sin\theta_W + A_\mu \cos\theta_W\,. \tag{206}$$

While W_3 couples only to the left-handed fermions, B couples with both ψ_L and ψ_R (albeit with different strength), and so Z, which has a small admixture of B, couples to the right-handed fermions too.

While the gauge boson A is massless, we are yet to show that it is indeed the photon. For this, we need to show that (i) its coupling with fermions is vectorial, *i.e.*, there is no γ_5-dependence, and (ii) the coupling strength is proportional to the electric charge Q.

Q. Given $M_W = 80.41$ GeV and $M_Z = 91.14$ GeV, and the Fermi constant $G_F = 1.16637\times 10^{-5}$ GeV^{-2}, use $G_F/\sqrt{2} = g^2/8M_W^2$ to calculate g, v, g', $\sin^2\theta_W$. Also calculate $g^2\sin^2\theta_W/4\pi$. Is it close to the fine structure constant?

Q. Calculate the vertex factors for ZZh and Ahh.

Q. Convince yourself, without an explicit calculation, that the Zhh vertex factor must be zero; the vertex cannot exist. [*Hint: Z is spin-1 and h is spin-0; so the final state must be in* $l=1$ *to conserve angular momentum. That gives rise to an antisymmetric wave function for the hh system. Now apply Bose statistics.*]

5.7 Recovering the photon

We are yet to discuss the mass generation mechanism for the fermions, so let us take them to be massless. The Dirac Lagrangian, in presence of the gauge fields, can be written as

$$\mathcal{L} = \bar\psi(iD_\mu\gamma^\mu)\psi = \bar\psi_L(iD_{L\mu}\gamma^\mu)\psi_L + \bar\psi_R(iD_{R\mu}\gamma^\mu)\psi_R\,. \tag{207}$$

Note that D_L and D_R are not the same:

$$\begin{aligned} D_L &= \partial + i\frac{g}{2}\vec\tau.\mathbf{W} + i\frac{g'}{2}Y_L B\,,\\ D_R &= \partial + i\frac{g'}{2}Y_R B\,. \end{aligned} \tag{208}$$

We can use $T_3 = \frac{1}{2}\tau_3$, $\frac{1}{2}Y_L = Q - T_3$, and $\frac{1}{2}Y_R = Q$, to write the gauge-fermion interaction as

$$\mathcal{L}_{f\bar fV} = -\bar\psi_L\left[gT_3W_{\mu 3} + g'(Q-T_3)B_\mu\right]\gamma^\mu\psi_L - \bar\psi_R\left[g'QB_\mu\right]\gamma^\mu\psi_R\,. \tag{209}$$

Using (206), one can write the coupling with A_μ as

$$\bar\psi_L\left[gT_3\sin\theta_W + g'(Q-T_3)\cos\theta_W\right]\gamma^\mu\psi_L + \bar\psi_R\left[g'Q\cos\theta_W\right]\gamma^\mu\psi_R\,. \tag{210}$$

From (202), $g\sin\theta_W = g'\cos\theta_W$. Let us call this e, the unit electric charge, so that both ψ_L and ψ_R couple to A_μ with a strength of eQ. Thus, A is indeed the photon field.

Q. Show that Z couples with the left- and right-handed fermions with coupling strengths of

$$\frac{g}{\cos\theta_W}\left[T_3 - Q\sin^2\theta_W\right]\,, \quad -\frac{g}{\cos\theta_W}Q\sin^2\theta_W\,, \tag{211}$$

respectively. Hence, show that if the $\bar{f}fZ$ coupling is written as

$$\mathcal{L}_{f\bar{f}Z} = -\frac{g}{2\cos\theta_W}\bar{f}\gamma^\mu(c_V - c_A\gamma_5)fZ_\mu\,, \tag{212}$$

then

$$c_V = T_3 - 2Q\sin^2\theta_W\,, \quad c_A = T_3\,. \tag{213}$$

Hence show that for electrons $c_V \ll c_A$. One may also define

$$c_L = c_V + c_A = 2(T_3 - Q\sin^2\theta_W)\,, \quad c_R = c_V - c_A = -2Q\sin^2\theta_W\,, \tag{214}$$

so that $c_L^2 + c_R^2 = 2(c_V^2 + c_A^2)$, and

$$c_V - c_A\gamma_5 = c_L P_L + c_R P_R\,. \tag{215}$$

5.8 Fermion masses

With ψ_L being a doublet and ψ_R a singlet under SU(2)$_L$, a mass term of the form $-m\bar{\psi}_L\psi_R$ is not gauge invariant. But again the Higgs field comes to our rescue: one can construct a gauge singlet from the combination $\bar{\psi}_L\psi_R\Phi$ [15]. The term has mass dimension 4 and hence to preserve renormalisability, the coupling should be dimensionless. This is known as a Yukawa term. Consider the interaction Lagrangian

$$\mathcal{L}_Y = -h_e\bar{L}_L e_R\Phi + h.c. \tag{216}$$

where L is the lepton doublet: $\begin{pmatrix}\nu_e\\ e\end{pmatrix}$. Shifting the vacuum, one gets a term like

$$\mathcal{L}_Y = -\frac{h_e v}{\sqrt{2}}\bar{e}_L e_R + h.c. \tag{217}$$

which gives the electron mass:

$$m_e = \frac{1}{\sqrt{2}}h_e v\,. \tag{218}$$

Thus the electron mass can be generated in a gauge invariant way. From the known value of $m_e = 0.511$ MeV, one obtains $h_e \approx 2.94\times 10^{-6}$. One can similarly generate μ and τ masses, and also the down-type quark masses by replacing L with the quark doublet Q and e with the quark singlet d.

Unfortunately, this simple prescription cannot generate masses for the $T_3 = +\frac{1}{2}$ fermions. The neutrinos are taken to be massless in the SM, so we need not bother, but there are the up-type quarks, in particular the top quark whose mass is about 175 GeV. We construct another doublet

$$\tilde{\Phi} = i\sigma_2\Phi^* = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\phi^-\\ \phi_0 - i\phi_3\end{pmatrix} = \begin{pmatrix}\phi_0 - i\phi_3\\ -\phi^-\end{pmatrix}\,, \tag{219}$$

---

[15] There are two doublets, from which you pick up the singlet combination, and multiply this with the singlet fermion, so that the result is a gauge singlet.

and the term $-h_u \bar{Q}_L u_R \tilde{\Phi} + h.c.$, after vacuum shifting, gives the mass term for the up quark:

$$m_u = \frac{1}{\sqrt{2}} h_u v \,. \tag{220}$$

Even if the neutrinos are taken to be massless, there are *nine* independent Yukawa couplings in the SM, six for the quarks and three for the leptons. There is a large hierarchy too; h_e, as we have just seen, is of the order of 10^{-6}, while h_t, the top Yukawa, is close to 1. Why this hierarchy? This is one of the puzzles of the SM.

Q. Show that the weak hypercharge of $\tilde{\Phi}$ is -1.
Q. Show that both $-h_e \bar{L}_L e_R \Phi$ and $-h_u \bar{Q}_L u_R \tilde{\Phi}$ have $Y = 0$.
Q. In the calculation of $e^+e^- \to \mu^+\mu^-$, why don't we consider the s-channel scalar exchange?

6 The Higgs Boson

The Higgs mechanism, discussed in the last chapter, must not be confused with the Higgs boson. The former is a mechanism to generate masses of gauge bosons in a gauge-invariant way. The latter is the remnant of the original complex doublet Φ in the spectrum, after all the Goldstones have been eaten up by the weak gauge bosons. In the literature, Φ is often called the Higgs doublet, but be sure that only one of its four components is the Higgs boson.

The vacuum shift generates the mass for all massive particles. It generates a mass $m_f = h_f v/\sqrt{2}$ for the fermions, $\frac{1}{2}gv$ for $W^\pm$ (for Z we multiply this by $\sec\theta_W$), and the Higgs boson itself gets a mass of $\sqrt{2\lambda v^2}$. For $m_h = 125$ GeV, $\lambda \approx 0.13$. (We know $v \approx 246$ GeV from the measured values of $m_W = 80.41$ GeV, α, and $\sin\theta_W$. Show how.)

Even before its discovery, there were some reasons to expect that the Higgs boson cannot be much more massive than a few hundreds of GeV. They come from the precision measurement of the electroweak observables, and also from the demand that the theory should remain well-behaved till at least the SM is valid and no new theory takes over. For example, if new physics — perhaps quantum gravity — appears at 10^{19} GeV, the Planck scale, we expect the Higgs potential to remain sensible till that point. Let us see what this rather enigmatic statement means, although such bounds might only be of academic interest now for the SM Higgs (but not if there are more than one such scalars).

6.1 Theoretical bounds on the Higgs boson mass

The theoretical upper bound on a SM-like scalar, which we will call the Higgs, come from two different sources: (i) Unitarity of gauge boson scattering, and (ii) Triviality of the quartic coupling. There are some further constraints on the Higgs sector too. While we cannot go into any detailed discussion of these topics, let us just see what they mean and how they act. Note that if there are more than one scalar in the theory, such constraints are individually valid on all of them.

6.1.1 Unitarity of gauge boson scattering

Consider the scattering $W^+W^- \to W^+W^-$. Without the Higgs, the scattering proceeds through the three- and four-point gauge vertices of the field strength tensor, and it can be shown that the amplitude grows rapidly with $\sqrt{s}$. With the introduction of the Higgs, there is a new three-point scalar-gauge-gauge vertex that softens this behaviour. Let us consider where we take all gauge bosons in their longitudinal mode. This is particularly instructive since the longitudinal modes are nothing but the Goldstones, which were eaten up. To a first approximation, if $m_W \ll \sqrt{s}$, we can write

$$\mathcal{A}(W_L^+ W_L^- \to W_L^+ W_L^-) = \mathcal{A}(w^+ w^- \to w^+ w^-) + \cdots\,, \tag{221}$$

where w's are the corresponding Goldstones. The neglected terms on the right-hand side are of the order of m_W^2/s or higher. The result is known as the *equivalence theorem.* The scattering will have two diagrams, one in s-channel and another in t-channel, and a fairly simple computation gives

$$\mathcal{A}(w^+ w^- \to w^+ w^-) = -\frac{m_h^2}{v^2}\left(\frac{s}{s - m_h^2} + \frac{t}{t - m_h^2}\right)\,. \tag{222}$$

We can also write, using the partial wave decomposition, $\mathcal{A}$ as

$$\mathcal{A} = 16\pi \sum_{l=0}^{\infty} (2l+1) P_l(\cos\theta) a_l\,, \tag{223}$$

where a_l is the spin-l partial wave and P_l's are the Legendre polynomials. The scattering cross-section is

$$\sigma = \frac{16\pi}{s}\sum_{l=0}^{\infty}(2l+1)|a_l|^2 = \frac{1}{s}\mathrm{Im}[\mathcal{A}_{\theta=0}]\,, \tag{224}$$

where we have used the optical theorem in the last step. But $P_l(1) = 1$ for all values of l, and hence

$$|a_l|^2 = \mathrm{Re}(a_l)^2 + \mathrm{Im}(a_l)^2 = \mathrm{Im}(a_l)\,, \tag{225}$$

which gives the constraint

$$|\mathrm{Re}(a_l)| < \frac{1}{2}\,. \tag{226}$$

In the $l = 0$ partial wave, one finds

$$a_0 = \frac{1}{16\pi s}\int_{-s}^{0} \mathcal{A}\, dt = -\frac{m_h^2}{16\pi v^2}\left[2 + \frac{m_h^2}{s - m_h^2} - \frac{m_h^2}{s}\log\left(1 + \frac{s}{m_h^2}\right)\right]\,, \tag{227}$$

and at the limit $s \gg m_h^2$, we get

$$a_0 = -\frac{m_h^2}{8\pi v^2}\,, \tag{228}$$

which gives the unitarity constraint $m_h \le 870$ GeV. If you consider the coupled channel $W_L^+ W_L^- \to (2W_L^+ W_L^- + Z_L Z_L)$ and its equivalent amplitude $w^+ w^- \to (2w^+ w^- + zz)$, the bound is even stronger:

$$m_h < 710 \text{ GeV}\,. \tag{229}$$

So, even before the discovery, we could have said that the Higgs cannot be more massive than 700 GeV.

Now suppose there is no Higgs boson. When should we be able to see some abnormal behaviour, like the violation of unitarity, in gauge boson scattering? Consider the limit $m_h^2 \gg s$ and expand the logarithm in (227) to get

$$a_0(w^+ w^- \to w^+ w^-) = -\frac{s}{32\pi v^2}\,. \tag{230}$$

But this must be less than $\frac{1}{2}$, so

$$\sqrt{s} < 1.8 \text{ TeV} . \tag{231}$$

This bound can be lowered further using more constrained channels:

$$\sqrt{s} < 1.2 \text{ TeV} . \tag{232}$$

Thus, even if there is no Higgs boson, we are bound to get some anomalous signatures from gauge boson scattering at the ballpark of 1 TeV. The reason why this bound is different from the earlier bound of 710 GeV is that the earlier bound answers the question: if a Higgs boson is there, how massive it can be? The latter bound answers a different question: if the Higgs boson is not there, when can we expect to see some new physics, some abnormal behaviour of the scattering partial waves?

6.1.2 Triviality

You have possibly seen in the course on QFT that the dimensionless couplings of a renormalisable theory depend on the energy scale at which they are measured. Although we have not shown explicitly that the Standard Model is renormalisable (and this is not an easy thing to do; Veltman and 't Hooft got the Nobel prize for that), let us take this fact as accepted; so we expect the quartic coupling $\lambda = \lambda(q^2)$ where q is the energy scale where we measure λ. The change of λ with q^2, also known as the renormalisation group (RG) equation, can be written as

$$16\pi^2 \frac{d\lambda}{dt} = 12\lambda^2 + 6h_t^2\lambda - \frac{1}{2}(3g'^2 + 9g^2)\lambda + \frac{3}{16}\left(g'^4 + 2g'^2g^2 + 3g^4\right) - 3h_t^4 + \cdots \tag{233}$$

where $t = \ln(q^2/q_0^2)$. q_0 is the scale where the value of the coupling is specified (the boundary condition needed to solve the differential equation), and q is the scale where we would like to find λ.

You should now be able to draw all the one-loop diagrams that contribute to the right-hand terms. For example, the first term comes from two h^4 vertices, taking all possible Wick contractions into account. The terms that are proportional to λ come from the diagrams with a h^4 vertex and a gauge or fermion loop on one of the legs. Note that only the top quark loop is numerically relevant. Similarly, the last term $-3h_t^4$ comes from a box diagram with top quarks, so that has four Yukawa couplings. The minus sign is due to the fermionic loop. Remember that the final state must be identical with the original four-point Green's function. Terms that are numerically insignificant have been lumped into the dots.

Now $m_h^2 \propto \lambda$ and so for large values of m_h, the first term dominates. Neglecting the other terms, we can solve the equation as

$$\lambda(q) = \frac{\lambda(q_0)}{1 - \frac{3}{4\pi^2}\lambda(q_0)\ln(q^2/q_0^2)} . \tag{234}$$

Now there is a chance that the denominator may become zero and $\lambda(q)$ may blow up. This singularity is known as a *Landau pole*. If you claim that no matter what q is, $\lambda(q)$ should never hit the Landau pole, that gives an upper bound on $\lambda(q_0)$ and hence on the Higgs boson mass:

$$m_h^2 < \frac{8\pi^2 v^2}{3\ln\left(\frac{\Lambda^2}{v^2}\right)} , \tag{235}$$

where we have taken $q_0 = v$, the electroweak symmetry breaking scale, and $q = \Lambda$, some high scale up to which the Standard Model might be valid [16]. If the upper scale is at infinity, then no matter what you take the value of $\lambda(q_0)$, it is bound to blow up, so the only sensible theory is where $\lambda(q_0) = 0$ to start with. Now, that is a non-interacting trivial theory, and hence the bound is known as the triviality bound. Of course, the upper bound cannot be infinity; it can at most be the Planck scale $\sim 10^{19}$ GeV, after which the quantum gravity effects must take over and the Standard Model is bound to be invalid.

6.1.3 Vacuum stability

On the other hand, if λ and hence m_h is small, the last term, which is proportional to h_t^4, dominates. This is negative, so this drives the coupling to smaller and smaller values. If we demand that $\lambda(q)$ cannot be negative whatever the value of q might be — if it does then the potential will be unbounded from below and will not have a stable vacuum — this gives a lower limit on the Higgs boson mass: If we demand that $\lambda(q)$ cannot be negative whatever the value of q might be — if it does then the potential will be unbounded from below and will not have a stable vacuum — this gives a lower limit on the Higgs boson mass:

$$\lambda(q) = \lambda(q_0) - \frac{3}{16\pi^2} h_t^4 \ln\left(\frac{q^2}{q_0^2}\right) \Longrightarrow m_h^2 v^2 > \frac{3}{2\pi^2} m_t^4 \ln\left(\frac{\Lambda^2}{v^2}\right) . \tag{236}$$

It is surprising that for $m_h = 125$ GeV, the electroweak vacuum remains stable almost up to the Planck scale. Recently, Degrassi *et al.* (arXiv:1205.6497) showed that the vacuum does not go negative below the Planck scale only if $m_h > 129.4 \pm 1.8$ GeV, but in a subsequent paper, Alekhin *et al.* (arXiv:1207.0980) showed that the uncertainty in the top mass plays a significant role here, and even $m_h = 125$ GeV might be allowed till the Planck scale.

Q. If $m_h = 125$ GeV, where is the Landau pole for the SM?
Q. Draw all the Feynman diagrams contributing to the right-hand side of eq. (233). Write down the fields and perform the Wick contractions. Get the numerical factors for λ^4, g^4, and h_t^4 terms.
Q. Suppose there is no new physics till $\Lambda = 10^{15}$ GeV, *i.e.*, the Standard Model remains valid. What is the upper bound on the Higgs mass? What should be the bound if new physics might be there at 10 TeV?
Q. Perform the same exercise for the vacuum stability bound coming from eq. (236).

6.2 Production of the Higgs boson

At an e^+e^- collider like LEP, the main production channel is through the *Bjorken process*: $e^+e^- \to Z^* \to Zh$. There is another channel, known as the vector boson fusion. For the Bjorken process, the intermediate Z^* is virtual because $m_h > m_Z$. The cross-section in the ultra-relativistic limit is

---

[16] There is a catch. Even if the coupling does not hit the Landau pole, it may become large, so large that the perturbation expansion breaks down. Remember that the RG β-function is nothing but a perturbation series; if we take two-loop diagrams, we will get terms like λ^3 and so on. Those higher-order terms can be neglected as long as perturbation is valid. That is why people quote two upper bounds: one where the coupling becomes infinite (the Landau pole), and one where the coupling ceases to be perturbative ($\lambda \geq 1$). Naturally, the latter gives a smaller number than the former.

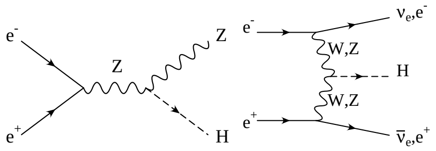

Figure 1: Higgs production processes at an e^+e^- collider.

given by (work this out as an exercise)

$$\sigma(e^+e^- \to Zh) = \frac{2\pi\alpha^2(c_L^2 + c_R^2)}{48(\sin\theta_W \cos\theta_W)^4} \frac{s + 12m_Z^2}{4(s - m_Z^2)^2}\,, \tag{237}$$

where c_L and c_R for electron can be worked out from (214). While working this out, do not forget to take into account the final state polarisation sum over Z. You will also need the ZZh vertex factor, given as an earlier exercise.

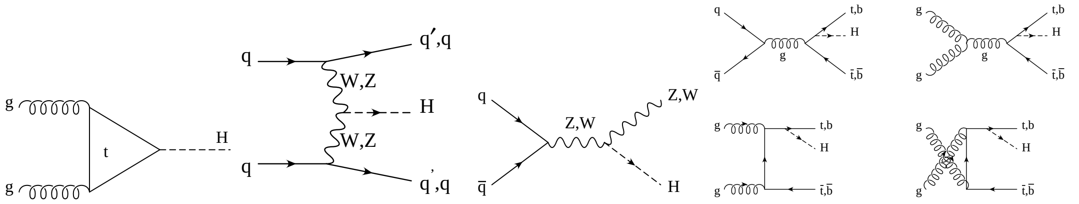

Figure 2: Leading Higgs production processes at hadron colliders: $gg \to h$, $qq \to qqh$, $q\bar{q} \to Wh, Zh$, and $qq \to t\bar{t}h$.

At the LHC, the dominant production is through the gluon-gluon fusion. Gluons, of course, do not couple to the Higgs directly, but this process occurs through a loop, as shown in Figure 2. The particle moving inside the loop is a top quark; that maximizes the Yukawa coupling and hence the amplitude. We can also have $q\bar{q} \to Z^* \to Zh$, but the sea antiquark content in the proton is not that high, and more significantly, the s-channel Z^* propagator gets a huge suppression due to the large value of $\sqrt{s}$.

The loop for gluon-gluon fusion appears to be divergent; there will be a d^4k integration, and there are only three fermion propagators that go like $1/k^3$. No diagram can be linearly divergent (*i.e.*, goes as the upper momentum cutoff Λ), because a linear term in momentum is not Lorentz-invariant. The diagram can at most be log-divergent (goes as $\ln\Lambda$); but even a cursory glance at the diagram tells us that this is bound to be finite. Why? Simply because there is no tree-level ggh term in the Lagrangian. If there were, we could have used the divergence coming from the loop as a counterterm to cancel the divergence in the tree-level coupling. If the tree-level coupling is absent, a renormalisable theory cannot have a divergence coming at one-loop level; there will be nothing to cancel the divergence with. Thus, the gluon fusion diagram is finite.

The cross-sections for various production channels at the LHC, running at $\sqrt{s} = 14$ TeV, is shown in Figure 3. The gluon-gluon fusion channel is clearly dominant; for a 125 GeV Higgs, the cross-section is about 10^4 fb. However, the predictions are not very well-defined; let us see why.

QCD uncertainties

Unlike LEP, Tevatron or LHC are hadronic machines. They collide hadrons, and hadrons mean

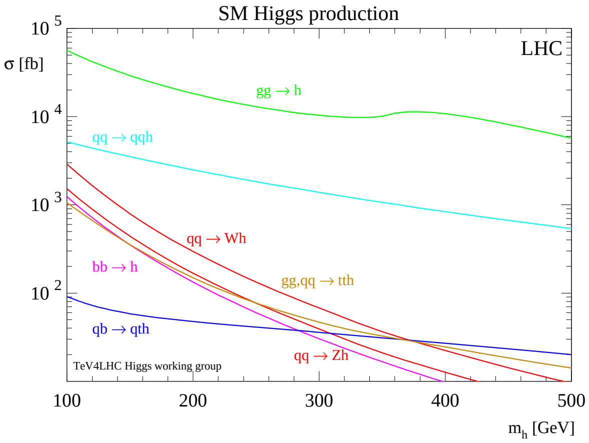

Figure 3: *Cross sections for SM Higgs boson production processes at the LHC* ($\sqrt{s}=14$ *TeV*).

strong interaction. Just like an e^+e^- collider will have QED corrections on the initial beams as well as the final state decay products, a hadronic collider will have QCD corrections. But the latter is much more important than the former, because the QCD gauge coupling $g_3 = \sqrt{4\pi\alpha_s}$ is more than 4 times larger than the QED gauge coupling $e = \sqrt{4\pi\alpha}$, so replacing even one external leg photon by a gluon increases the correction by a factor of 16, and then there are eight gluons! That is why it is imperative to calculate the higher-order QCD corrections as accurately as possible.

Let us consider the gluon-gluon fusion only, but this applies to any process in a hadronic collider. The leading-order (LO) cross-section for $pp \to h$, assuming $gg \to h$ is the only contributing process, is given by

$$\sigma(pp \to h) = \int dx_1 dx_2 f_g^p(x_1) f_g^p(x_2) \sigma(gg \to h)\,, \tag{238}$$

where f_g's are the parton distribution function (PDF) for the gluons. One can calculate the next higher order effects, mostly QCD corrections, on the cross-section. These effects are called next-to-leading-order (NLO) effects and shift the LO prediction. The exact amount of the shift depends on whether it is a pp or a $p\bar{p}$ collider, and also on its centre-of-mass energy $\sqrt{s}$. The ratio σ_{NLO}/σ_{LO} is called the K-factor at the NLO level:

$$K_{NLO} = \frac{\sigma_{NLO}}{\sigma_{LO}}\,, \tag{239}$$

and we must remember that the PDFs are different for LO and NLO. At LHC running at $\sqrt{s} = 7$ TeV, $K_{NLO} \approx 1.8$, so that is a rather significant correction. In fact, people have computed till next-to-next-to-leading order (NNLO) for the process $gg \to h$, and that gives an enhancement by about a factor of 2.3-2.5 over the LO cross-section. The figure shows the NNLO cross-sections (or NLO for those processes where an NNLO calculation is not yet available) for Higgs production.

One truncates at the NNLO not because the next-order results are negligible; hopefully they are, but they are yet to be computed. Even at NNLO, there are two major uncertainties, both contributing at the level of $\sim 20\%$. The first comes from the PDFs. There are many PDFs in the market (even at the NNLO level, there are at least three; remember that PDFs cannot be found from any first principle and are basically some model), and changing the PDFs changes the result by about 20%. Of course, one would ultimately know, when the Higgs is produced, which is the best PDF, but till then all are equally good.

6.3 Decay of the Higgs boson

At the tree-level, the Higgs boson can decay to two gauge bosons (W^+W^- or ZZ), or a fermion-antifermion pair $f\bar{f}$. The decay widths are given by

$$\begin{aligned}
\Gamma(h \to ZZ) &= \frac{G_F m_h^3 \beta_Z}{64\sqrt{2}\pi}\left(4 - 4a_Z + 3a_Z^2\right) , \\
\Gamma(h \to W^+W^-) &= \frac{G_F m_h^3 \beta_W}{32\sqrt{2}\pi}\left(4 - 4a_W + 3a_W^2\right) , \\
\Gamma(h \to f\bar{f}) &= \frac{G_F m_f^2 m_h N_c \beta_f^3}{4\sqrt{2}\pi} . \qquad (240)
\end{aligned}$$

Here we have used the standard notations: $\beta_x = 1 - 4m_x^2/m_h^2$ and $a_x = 1 - \beta_x^2 = 4m_x^2/m_h^2$. For leptons in the final state $N_c = 1$ while for quarks it is 3. The first two decays are not allowed for $m_h \approx 125$ GeV, but even below the $2W$ threshold the Higgs can decay to WW^*, followed by the $f\bar{f}$ decay of the virtual W. For the fermions, it is obvious that the decay width will be almost completely dominated by $b\bar{b}$.

Apart from these tree-level decays, h can decay to two photons, two gluons, or $Z\gamma$ through a fermion loop, similar in nature to the one we considered for the Higgs production by gluon-gluon fusion. Their contributions to the total decay width is, however, rather small, but the 2γ channel is a clean one to detect. The LHC experiments found a significant excess in the $h \to \gamma\gamma$ channel for $m_h = 125$ GeV. How significant? Roughly speaking, if there are S number of events and B number of expected background events (maybe estimated through a computer simulation), the ratio $S/\sqrt{S+B}$, which can be approximated by $S/\sqrt{B}$ if $B \gg S$, should be about 3 for a good signal and about 5 for a confirmed discovery. With 100 background events, we need at least 130 events to be confident that it is not a statistical fluctuation and indeed a signal, and with 150 events we can confidently announce the discovery. Both ATLAS and CMS experiments at the LHC got such a discovery signal for the Higgs, when all possible decay channels are combined. In other words, the chances that the background fluctuated to generate such an excess is about 0.00003% or even less.

Q. We have already observed the modes $h \to \gamma\gamma$ and $h \to ZZ$. From this, show that the spin of h cannot be 1. This is known as the *Landau-Yang theorem*.
Hint: Consider a spin-1 object with polarization $\mathbf{e}_V$ *decaying to two photons* $(\mathbf{e}_1, \mathbf{k})$ *and* $(\mathbf{e}_2, -\mathbf{k})$. *Gauge invariance tells us that* $\mathbf{e}_1.\mathbf{k} = \mathbf{e}_2.\mathbf{k} = 0$. *The amplitude can only depend on*

$$[(\mathbf{e}_1 \times \mathbf{e}_2).\mathbf{e}_V]\, F_1(k^2)\,,\ [(\mathbf{e}_1 \times \mathbf{e}_2).\mathbf{k}]\,(\mathbf{e}_V.\mathbf{k})F_2(k^2)\,,\ (\mathbf{e}_1.\mathbf{e}_2)(\mathbf{e}_V.\mathbf{k})F_3(k^2)\,. \qquad (241)$$

Now apply Bose statistics that $M(\mathbf{e}_1, \mathbf{e}_2, \mathbf{k}) = M(\mathbf{e}_2, \mathbf{e}_1, -\mathbf{k})$. *What does the last condition physically mean?*

6.4 Problems with the Higgs boson

The major problem with an elementary Higgs boson is the fine-tuning of its mass. This is characteristic of any $\lambda\phi^4$ theory; while the theory is renormalizable, the quantum correction to the Higgs mass squared is of the order of Λ^2 where Λ is the scale up to which the theory is valid. Assuming $\Lambda = 10^{19}$ GeV and the Higgs mass of the order of 100 GeV, this gives a fine-tuning at the order of 1 in 10^{34} between the bare mass term squared and the counterterm.

This is expected. After all, setting $m_h = 0$ does not restore any symmetry of the Lagrangian. So, the quantum correction is not protected by any symmetry. This is to be compared with gauge boson masses, protected by gauge symmetry (so that setting gauge boson masses equal to zero restores gauge symmetry) or fermion masses, protected by chiral symmetry. It can be shown that the mass corrections are of the order of $\log \Lambda$ for these two cases.

There are several ways to solve the fine-tuning problem.

1. One can invoke a symmetry between bosons and fermions. Suppose there is a theory in which there is a boson for every SM fermion and vice versa. If the corresponding couplings are also related, in the limit where this symmetry is exact, the quadratically divergent contribution to m_h^2 cancels — remember that fermion loops come with an extra minus sign. Such a symmetry is known as *supersymmetry.* Supersymmetry is definitely not exact; we have not seen any spin-0 partner of electrons at 0.5 MeV, in fact, we have not seen the supersymmetric partners for any SM particle. But even if it is broken, we expect the superparticles to be not too massive, so that the best feature of supersymmetry, *viz.* cancellation of the quadratic divergence, is not too spoilt. Thus, if superpartners exist at all, they should be accessible at the LHC.

   Supersymmetry is not a single model; rather, it is a conglomeration of a large class of models. In all these models, a general feature is that there is a low-mass scalar (there are more than one Higgs-like scalars in supersymmetry) which must be less than 135 GeV, so the discovered Higgs boson might be one of supersymmetry. But no superparticles have yet been discovered.

2. Another option is to make the cutoff scale of the theory low, say of the order of 10 TeV. For this, one needs to consider more than four space-time dimensions, but the extra dimensions are all small and curled up. We cannot go into any detail of such extra dimensional models; it is enough to say that they solve the fine-tuning problem by bringing down the effective Planck mass from 10^{19} GeV to a few TeV.

3. The third option is to consider the Higgs as a fermion-antifermion composite. If it is a composite object like mesons or the Cooper pair, the fine-tuning problem, which is a signature for elementary scalars, vanishes. This is the basic idea of Technicolour models. However, such models are severely constrained from experimental data and the simplest versions are ruled out.

4. One can even think of a bigger symmetry group, and the Higgs appearing as a Goldstone boson from spontaneous breakdown of some global symmetry. Goldstone bosons are massless, and that is why the Higgs mass is so small compared to the Planck scale; the observed mass is only a higher-order quantum effect. Such models are known as Little Higgs models.

5. One can just say that there is some unknown symmetry which makes the sum of all quadratically divergent corrections to the Higgs mass vanish. This translates to

   $$m_h^2 + 2m_W^2 + m_Z^2 - 4m_t^2 \approx 0\,, \tag{242}$$

   which is known as the Veltman condition. However, it is obvious that this condition is not satisfied in the Standard Model.

6.5 The Standard Model

We have assembled all pieces for the Standard Model, which we will call SM for brevity from now on. This is what we have.

- The gauge group is $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. Thus, it is not a unified gauge theory; there are three different coupling constants for the three gauge groups. The first group, $\mathrm{SU}(3)_c$, is an exact gauge theory of strong interaction, known as Quantum Chromodynamics (QCD). It is noteworthy that the SM includes strong interaction too and not just the electroweak part.
- The last two groups describe weak interaction and electromagnetism. It is where the mixing occurs. However, we still do not have a unified theory in the proper sense of the word — there are two different gauge couplings g and g'. There are two remarkable things that happen here.
  1. The two neutral gauge bosons W_3 and B mix in such a way that they reproduce the Z and a massless photon. Not only that, the photon couplings turn out to be vectorial in nature.
  2. The Higgs mechanism, which is in essential component of the electroweak unification, provides mass to the weak gauge bosons in a renormalizable way. It also makes the fermions massive.

Why do we still call it a model then and not a theory? This is mostly a case of semantics, and also historical legacy; you can call it the theory of particle interactions if you like. However, problems like fine-tuning (see below) make us believe that the SM is only a partial story, however successful it might be. After all, every theory has its region of validity, outside which it may not be valid. Newton's laws of gravity are not enough to describe the perihelion motion of mercury. Even General Relativity may not be enough when the energy scale becomes too high and the quantum aspects become important. Thus, it is no discredit that the SM might cease to be valid after some energy scale. The point is that the limit might be just too close; also, the theoretical drawbacks are quite glaring.

Is there any evidence to go beyond the SM? The existence of dark matter is one; there are about 5 times more dark matter in the universe than visible, *i.e.*, SM matter, but there is no dark matter candidate in the SM. One might also point at massive neutrinos; the SM neutrinos are massless. Remember that fermion masses come from the interaction of the Higgs field to both the right- and left-chiral fermions, $\bar{\psi}_L\psi_R\phi$. There are no ν_R to start with, so the neutrinos are massless. Strictly speaking, that is not a drawback; if it were known during the 1960s that neutrinos are massive, it would have been easy to incorporate that in the SM with three ν_R fields thrown in.

However, ν_R fields are strange objects in the sense that they do not have any gauge interactions; they are leptons, charge neutral, and right-chiral, so do not interact with any gauge bosons. Having such weird objects would be something new, and in that sense this might be something beyond the SM.

Again, it is not necessary to have ν_R fields. Neutrinos might very well be two-component spinors, but they might be identical with their antiparticles. Photons and neutral π mesons are examples of particles which are their own antiparticles. No fermion except for the neutrinos can ever dream to be its own antiparticle, because no fermion except the neutrinos are charge neutral, and antiparticles must have the opposite electric charge of the particle.

One can formulate a field theory of such fermions which are their own antiparticles. They are called *Majorana fermions*, after Ettore Majorana who first gave the theory. Obviously, if particles and antiparticles are the same, lepton number, and in general fermion number, will be violated. Technically, that is not a big problem. These are called discrete global symmetries and can be violated without any bad consequences (breaking of a continuous global symmetry would lead to unacceptable massless Goldstone bosons). If neutrinos are Majorana fermions, one can generate the mass with only the left-chiral fields.

There is a third problem, known as the baryon asymmetry of the universe. The baryon number of the universe is not zero, we see a lot of baryons but no antibaryons in nature. As far as we know, there are no antimatter galaxies. Thus, there is a net baryon excess, given by the ratio of baryon and photon numbers of the universe: $n_b/n_\gamma \sim 10^{-9}$. However, within the SM, you can generate an asymmetry only at the level of 10^{-18}.

7 Flavour Physics

Let us start this section with two puzzles.

- One can measure the Fermi constant G_F from different weak decays. If the weak interaction is a single force, all measurements should yield the same value of G_F. One can measure G_F very precisely from muon decay $\mu^- \to e^-\nu_\mu\bar{\nu}_e$, where the muon lifetime $\tau_\mu = 1/\Gamma_\mu$ is given by

  $$\frac{1}{\tau_\mu} = \frac{G_F^2 m_\mu^5}{192\pi^3} F\left(\frac{m_e^2}{m_\mu^2}\right)\left(1 + \frac{3}{5}\frac{m_\mu^2}{m_W^2}\right)[1+\xi] \,, \tag{243}$$

  where the phase space factor

  $$F(x) = 1 + 8x + 8x^3 + x^4 + 12x^2 \ln x \,, \tag{244}$$

  and ξ is a small number which includes the electroweak radiative corrections. The number turns out to be $G_F = 1.166367 \times 10^{-5}$ GeV^{-2}. However, G_F obtained from neutron β-decay is smaller than this value: $G_F^2(n) \approx 0.96 G_F^2(\mu)$ [17].

  While one can try to attribute this discrepancy to the structural difference of neutron and proton, the discrepancy is much more serious in the semileptonic K-meson decay: $K^+ \to \pi^0\mu^+\nu_\mu$. This is, again, a weak decay, but $G_F^2(K^+) \approx 0.04 G_F^2(\mu)$. Why?

- Consider the two possible K-meson decays: $K^+ \to \pi^0\mu^+\nu_\mu$ and $K^+ \to \pi^+\nu\bar{\nu}$. Both are weak processes (otherwise the strange quark cannot decay), but the quark-level subprocesses are $s \to u\mu\bar{\nu}_\mu$ for the first case and $s \to d\nu\bar{\nu}$ for the second case. The first one proceeds through the exchange of a virtual W and is known as a charged-current (CC) process. The second one, similarly, should proceed through a tree-level Z exchange (photon cannot change the flavour) and is known as a flavour-changing neutral current (FCNC) process. Surprisingly, while the CC decay has a large decay width (3.35% of the K^+ decay through this channel), the FCNC decay is almost absent in the SM (only about 1.7×10^{-8}% K^+, or about 1 in 10^{10}, has this decay)! This is true for other FCNC processes too; all are either absent or terribly suppressed compared to expectation. Again, why?

---

[17] The neutron half-life is much larger than the half-life of muon, but that is entirely due to the very narrow phase space for the former decay, resulting from $m_n - m_p \ll m_n$.

7.1 The Cabibbo mechanism

Consider only three quarks u, d, and s, that existed in Gell-Mann's time. A central idea of the quark model is that there are strange mesons, like the K, that decay into the nonstrange mesons through weak interaction, and one postulates that the weak interaction violates strangeness. Let us see how this can come about.

The d quark decays to the u quark through the charged-current weak interaction:

$$\mathcal{L}_{cc} = -\frac{g}{\sqrt{2}}\left[\bar{u}\gamma^\mu P_L d W^+_\mu + \text{h.c.}\right]\,, \tag{245}$$

and this is responsible for the β-decay of neutron. Thus, the charged weak current (CC) changes flavour within a generation. However s does not yet have a partner; how can it decay? The answer is that it must decay to, again, the u quark, by some mechanism that does not remain confined within a generation. Suppose, as a first step, we predict a similar form of $\mathcal{L}_{cc}$ for the $s \to u$ transition.

We have seen earlier that pinching off the W propagator (reducing the two weak vertices to a single contact four-Fermi interaction) results in the celebrated formula

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2}\,. \tag{246}$$

If we assume the same g for weak interaction, the value of G_F obtained from β-decay, μ decay, or K decay, should be the same. We have just seen that they are not. Instead of predicting a new gauge coupling for every weak decay, let us just modify (245):

$$\begin{aligned}\mathcal{L}^d_{cc} &= -\frac{g\cos\theta}{\sqrt{2}}\left[\bar{u}\gamma^\mu P_L d W^+_\mu + \text{h.c.}\right]\,,\\ \mathcal{L}^s_{cc} &= -\frac{g\sin\theta}{\sqrt{2}}\left[\bar{u}\gamma^\mu P_L s W^+_\mu + \text{h.c.}\right]\,.\end{aligned} \tag{247}$$

Comparing with the experiments, we see that $\cos^2\theta \approx 0.96$, $\sin^2\theta \approx 0.04$ (and it is a good thing that they add up to one). With this in mind, this is what Nicola Cabibbo proposed:

- There are two relevant bases for the two down-type quarks. One is where the u quark couples only with the d quark and not to the s. In this basis, weak interaction does not allow crossover between generations. We will call this the *flavour basis*; this is also called, in various texts, the *gauge basis* or the *weak basis*. This is not a stationary basis; the states are not physical states or mass eigenstates. If they were mass eigenstates, the mass term would have been like

  $$\mathcal{L}_{mass} = -m_d\bar{d}d - m_s\bar{s}s \tag{248}$$

  but no cross-terms of the form $m_{ds}\bar{d}s$. In the flavour basis, such cross-terms are in general present.

- If these flavour eigenstates are not mass eigenstates, one must rotate the basis and get to the *mass basis*, consisting of the stationary states. These are, then, the physical quarks that we observe (well, free quarks are not observable, so let us just say that this is the physical basis). To avoid confusion, let us denote the flavour basis by a prime (d', s') and the mass basis without a prime (d, s). There is only one charge-$\frac{2}{3}$ quark; so there is no distinction between u and u'. While this scenario will change once we have more quarks, right now let us stick to the three-quark model.

If we apply a rotation

$$d' = d\cos\theta + s\sin\theta\,, \quad s' = s\cos\theta - d\sin\theta\,, \tag{249}$$

we see that (247) is precisely restored from (245), if the former is written in the mass basis and the latter in the gauge basis. The angle θ is known as the *Cabibbo angle*; experimentally, $\cos^2\theta \approx 0.96$.

7.2 Flavour-changing neutral current

Right now, let us neglect the mixing between W_3 and B and assume $Z = W_3$, *i.e.*, we put the Weinberg angle $\theta_W = 0$. This does not affect what we are going to discuss here, and the rest of this chapter. If W^+, W^- and Z form a triplet, there must be a neutral-current (NC) partner of (245):

$$\mathcal{L}_{nc} = -\frac{g}{\sqrt{2}}\left[\bar{u'}\gamma^\mu P_L u' Z_\mu + \bar{d'}\gamma^\mu P_L d' Z_\mu\right]\,. \tag{250}$$

There is no such term involving the s' quark, because that is, at least in the 3-quark model, a singlet, and does not have weak interaction; it is only through the mixing with d' that it decays. In the mass basis, the term with up-quark current does not change, but the second term changes:

$$\mathcal{L}_{nc} = -\frac{g}{\sqrt{2}}\left[(\bar{d}\cos\theta + \bar{s}\sin\theta)\gamma^\mu P_L(d\cos\theta + s\sin\theta)Z_\mu\right]\,. \tag{251}$$

There are four terms coming from here; two of them involve flavour-conserving neutral currents ($\bar{d}\cdots d$ and $\bar{s}\cdots s$ terms), but the rest two are more disturbing:

$$\mathcal{L}_{fcnc} = -\frac{g\sin\theta\cos\theta}{\sqrt{2}}\left[(\bar{d}\gamma^\mu P_L s + \bar{s}\gamma^\mu P_L d)Z_\mu\right]\,. \tag{252}$$

This introduces severe flavour-changing neutral currents (FCNC) in the theory, untenable by experiments [18]. Somehow, we need to eliminate these terms without abandoning the good features of the Cabibbo theory.

7.3 The Glashow-Iliopoulos-Maiani (GIM) mechanism

The idea of Glashow, Iliopoulos, and Maiani (GIM) to deal with the FCNC problem has the dual quality of elegance and simplicity, and at that time it was a revolutionary one. Suppose that the s quark is not alone; it has a partner, which let us call charm (c), which is a charge-$\frac{2}{3}$ quark, and thus $\begin{pmatrix} c \\ s \end{pmatrix}$ forms an identical doublet to the $\begin{pmatrix} u \\ d \end{pmatrix}$. The only reason that charm was not discovered at that time was its heaviness; in fact, the GIM group calculated its mass to be close to 1.5 GeV, a surprisingly accurate prediction, from the phenomenon of K^0-$\overline{K}^0$ mixing, which we will discuss later in this chapter.

What happens if there is a charm quark? The strange quark feels the weak interaction too, because it is a part of the second doublet. Thus, we can modify (250) as

$$\mathcal{L}_{nc} = -\frac{g}{\sqrt{2}}\sum_{i=u',d',s',c'}\left[\bar{q'_i}\gamma^\mu P_L q'_i Z_\mu\right]\,. \tag{253}$$

---

[18] It is easy to see that the decay width of $K^+ \to \pi\nu\bar{\nu}$ will be larger than that of $K^+ \to \pi^0\mu^+\nu_\mu$ by a factor of $\cos^2\theta/\sin^2\theta \approx 24$, and if you take all three generations of neutrinos into consideration, that will be 72.

Now, of course, the Hilbert space of charge-$\frac{2}{3}$ quarks is a two-dimensional one, and one should in principle rotate the (u', c') states too to the mass basis of (u, c). However, as we will see, the neutral current part does not depend on the magnitude of this rotation, and the charged current strength is proportional to the relative misalignment of the (u, c) and the (d, s) basis. Thus, without the loss of any generality, we can keep one fixed and rotate the other; in other words, we can say that the charge-$\frac{2}{3}$ quarks were stationary states to begin with, and we have to rotate the charge-$-\frac{1}{3}$ basis. One can do the opposite but the physics does not change.

So suppose we take $u' = u$ and $c' = c$, but write d' and s' in terms of d and s. It is a simple exercise to show that

$$\mathcal{L}_{nc} = -\frac{g}{\sqrt{2}}\left[\bar{d}_i\gamma^\mu P_L d_i Z_\mu + \bar{s}_i\gamma^\mu P_L s_i Z_\mu\right] . \tag{254}$$

Thus, there is no FCNC, and we have achieved what we are trying to do. A bonus is the universality of the flavour-conserving NC interactions; they come with the same strength. We also note that (247) does not change but there is a similar term involving the charm quark; $c \to s$ strength is proportional to $\cos\theta$ and $c \to d$ strength to $\sin\theta$. The angle θ measures the relative misalignment between the (u, c) and the (d, s) bases.

The mixing matrix has been parametrised by only one element θ. Any such transformation matrix is unitary, and is one element enough to parametrise a 2×2 unitary matrix? To answer this question, let us perform a general exercise of counting the number of independent elements in an $N \times N$ quark mixing matrix, *i.e.*, the mixing matrix if there were N such quark generations.

To start with, any $N \times N$ unitary matrix has N^2 entries and hence $2N^2$ real elements. But not all of them are independent. The constraint condition is $UU^\dagger = 1$, which translates to $U_{ij}U^*_{kj} = \delta_{ik}$. If $i = k$, there are N such equations with 1 on the right-hand side, and the left-hand side is by construction real. If $i \neq k$, there are ${}^NC_2 = \frac{1}{2}N(N-1)$ such conditions, but each equation is actually a sum of two equations, for the real and the imaginary parts, both of which must be equated to zero. Thus, there are $N + 2 \times \frac{1}{2}N(N-1) = N^2$ constraints and there are only N^2 number of independent elements.

However, this mixing matrix is sandwiched between two quark wavefunctions, and one can redefine and absorb a phase in any wavefunction. This further reduces the number of independent parameters, not by constraint equations but by absorbing them elsewhere. Consider the 2×2 mixing matrix, with which the CC Lagrangian looks like

$$\begin{pmatrix} \bar{u} & \bar{c} \end{pmatrix} \begin{pmatrix} Ae^{i\alpha} & Be^{i\beta} \\ Ce^{i\gamma} & De^{i\delta} \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} , \tag{255}$$

where we have dropped all other constants, Dirac matrices, and the gauge fields, just for brevity. The constants $A, B, C, D, \alpha, \beta, \gamma$ and δ are related by 4 constraint equations. Now redefine

$$\tilde{u} = ue^{-i\alpha} , \quad \tilde{c} = ce^{-i\gamma} , \tag{256}$$

so that the CC Lagrangian becomes

$$\begin{pmatrix} \bar{\tilde{u}} & \bar{\tilde{c}} \end{pmatrix} \begin{pmatrix} A & Be^{i(\beta-\alpha)} \\ C & De^{i(\delta-\gamma)} \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} . \tag{257}$$

Now with a further redefinition $\tilde{s} = se^{i(\beta-\alpha)}$, we can absorb the phase with B, but that is the most that we can do; we have no further freedom to absorb the phase of D because both $\tilde{c}$ and $\tilde{s}$ fields are defined. Similarly, for an $N \times N$ matrix, we can absorb, at the most, $2N - 1$ phases. That gives the number of independent parameters to parametrise the quark mixing matrix as

$N^2 - (2N-1) = (N-1)^2$; so we need one parameter for two generations and four for three generations.

Are they all real parameters? To answer that question, let us just redo this exercise with orthogonal $N \times N$ matrices. There are N^2 elements to start with. From the equation $OO^T = 1$, we get N equations of the form $\sum_j |O_{ij}|^2 = 1$ and ${}^N C_2 = \frac{1}{2}N(N-1)$ equations of the form $O_{ij}O_{kj} = 0$ with $i \neq k$. The elements are real and thus there is no chance of phase absorption with redefinition of fields, and the number of independent elements is just $N^2 - N - \frac{1}{2}N(N-1) = \frac{1}{2}N(N-1)$.

Thus, among the $(N-1)^2$ elements of the unitary matrix, $\frac{1}{2}N(N-1)$ elements are real, and the rest, $\frac{1}{2}(N-1)(N-2)$ are phases. For two generations, the only independent element is real; remember the structure of the SO(2) group. For three generations, there are three real angles and one complex phase. Now, what are we going to do with that phase?

7.4 The K meson system

Let us first briefly recapitulate the role the K-mesons played in particle physics. They were first discovered in cosmic-ray experiments, as unstable particles that decay into pions. Neutral mesons do not leave a track but their characteristics, like mass and lifetime, can be guessed by studying the charged decay products that leave their characteristic tracks. For example, if a neutral K meson decays into a π^+ and a π^-, one can observe these π-meson tracks, measure their momenta, and calculate the invariant mass $M^2 = (p_{\pi^+} + p_{\pi^-})^2$ which must be the mass of the parent K meson. One can also estimate the lifetime of this invisible K meson by measuring the distance between the vertex where it was created (a characteristic kink) and the vertex where it decayed.

It turned out that there are two such objects that do not fall into the known particle category (it was in the 1940s). One of them decays to two pions, and the other to three pions. Their masses are identical, so it is tempting to identify them as the same particle, but the decay pattern is strange (that is why they were called strange mesons, and the name stuck to the quark too). We have seen earlier that parity of pion is negative, which can be ascribed to the opposite intrinsic parities of fermions and antifermions. So

$$P|2\pi\rangle = |2\pi\rangle\,, \quad P|3\pi\rangle = -|3\pi\rangle\,, \tag{258}$$

as parity is a multiplicative quantum number. How can the same particle decay into two different parity states if parity is conserved?

To answer this question, Tsang-Dao Lee and Chen-Ning Yang, two Chinese American physicists, conjectured that parity may not be conserved after all in weak interactions [19]. We have already seen how Wu *et al.* and Goldhaber *et al.* showed that weak interaction is purely left chiral in nature and hence violates parity maximally. So, it is possible for the same particle to decay into two different parity states through weak interaction.

That the weak interaction has a $V-A$ form was shown by Feynman and Gell-Mann, and simultaneously by Marshak and Sudarshan. Now it is easy to understand why the weak interaction should conserve the combined operation CP. Consider eq. (245). Under CP conjugation,

$$\bar{u}\gamma^\mu(1-\gamma_5)dW^+_\mu \longrightarrow \bar{d}\gamma^\mu(1-\gamma_5)uW^-_\mu\,, \tag{259}$$

---

[19] According to Feynman, it was an experimentalist, Martin Bloch, who first raised this idea in a conference, but Lee and Yang concretised the idea, and the credit is certainly due to them.

which is nothing but the hermitian conjugate (you have already done the necessary algebra when you went through the transformation of Dirac bilinears under P and C separately). Thus, if we start with eq. (245) with the hermitian conjugate included, under CP conjugation we get an identical equation, so the weak interaction must be invariant under CP. There is a subtle catch, which we will come to later.

Now consider the two-level system of K^0 and $\overline{K}^0$. One can write

$$CP|K^0\rangle = -|\overline{K}^0\rangle\,, \quad CP|\overline{K}^0\rangle = -|K^0\rangle\,, \tag{260}$$

where the minus sign comes from the parity eigenvalue, but this is a convention and one can put any phase η here with $|\eta|^2 = 1$. With this convention, one can define the CP eigenstates

$$|K_1\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle + |\overline{K}^0\rangle\right)\,, \quad |K_2\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle - |\overline{K}^0\rangle\right)\,, \tag{261}$$

so that

$$CP|K_1\rangle = -|K_1\rangle\,, \quad CP|K_2\rangle = |K_2\rangle\,. \tag{262}$$

Thus K_1 and K_2 are CP eigenstates with eigenvalues -1 and $+1$ respectively. We also have

$$CP|2\pi\rangle = |2\pi\rangle\,, \quad CP|3\pi\rangle = -|3\pi\rangle\,, \tag{263}$$

where 2π is a shorthand for $\pi^+\pi^-$ or $\pi^0\pi^0$, and 3π is a shorthand for $\pi^+\pi^-\pi^0$ or $\pi^0\pi^0\pi^0$ states. If weak interaction conserves CP, K_1 must go to 3π and K_2 must go to 2π. As the available phase space is much smaller for the 3π final state than the 2π final state, we expect K_1 to have a much longer lifetime than K_2, which is experimentally verified. The physical state K_L that has a longer lifetime (L for large/long) can thus be identified with K_1 and the physical state K_S with a shorter lifetime (S for small/short) can be identified with K_2.

7.5 Digression: two-level quantum mechanics

For a one-dimensional Hilbert space, one can write the wave equation as

$$H\psi = i\frac{\partial\psi}{\partial t}\,, \tag{264}$$

and the evolution is given by $\psi(t) = \exp(-iHt)\psi(0)$. Now consider a two-dimensional Hilbert space. If we are clever enough to choose the stationary basis, *i.e.*, the basis where the Hamiltonian is diagonal and the states are physically observable states, the wave equation is analogous to (264), but with an added index that can take the values 1 and 2:

$$H_i\psi_i = i\frac{\partial\psi_i}{\partial t}\,. \tag{265}$$

Now suppose we are not that clever, or the physics demands that we start with a non-stationary basis (this is what happens for Stark effect or the ammonia molecule, if you remember). In this basis, H is not diagonal, the states are not physical states, and there is a nonzero transition probability between the two states of the basis, which is given by the off-diagonal elements H_{12} and H_{21} of the Hamiltonian. We accordingly modify (265) to

$$H_{ij}\psi_j = i\frac{\partial\psi_i}{\partial t}\,. \tag{266}$$

This is a set of coupled differential equations. The best way to proceed is to diagonalise H and get the stationary states. Note that the hermiticity of H demands H_{11} and H_{22} to be real and $H_{12} = H_{21}^*$. Let us write

$$H = \begin{pmatrix} a & b \\ b & a \end{pmatrix} , \tag{267}$$

where, without loss of any generality, we have taken b to be real too. The trace is $2a$ and the determinant is $a^2 - b^2$, so the eigenvalues are $a + b$ and $a - b$ respectively, for the stationary states $|1\rangle$ and $|2\rangle$, defined as

$$|1\rangle = \frac{1}{\sqrt{2}} \left[|I\rangle + |II\rangle \right] , \quad |2\rangle = \frac{1}{\sqrt{2}} \left[|I\rangle - |II\rangle \right] , \tag{268}$$

where $|I\rangle$ and $|II\rangle$ are the basis states for the initial non-stationary basis. This can be written the other way around:

$$|I\rangle = \frac{1}{\sqrt{2}} \left[|1\rangle + |2\rangle \right] , \quad |II\rangle = \frac{1}{\sqrt{2}} \left[|1\rangle - |2\rangle \right] . \tag{269}$$

The time-dependence of $|1\rangle$ and $|2\rangle$ are given by

$$|1(t)\rangle = \exp[-i(a+b)t] \; |1(0)\rangle , \quad |2(t)\rangle = \exp[-i(a-b)t] \; |2(0)\rangle . \tag{270}$$

Now suppose at $t = 0$ the wavefunction was a pure $|I\rangle$. How should it evolve? Well, this is not a stationary state, but it can be written as a combination of $|1\rangle$ and $|2\rangle$, and we need to consider their evolutions, so

$$\begin{aligned} \psi(t) &= \frac{1}{\sqrt{2}} \left(|1(t)\rangle + |2(t)\rangle \right) \\ &= \frac{1}{\sqrt{2}} \left[e^{-i(a+b)t} |1(0)\rangle + e^{-i(a-b)t} |2(0)\rangle \right] \\ &= \frac{1}{2} \left[e^{-i(a+b)t} \left\{ |I\rangle + |II\rangle \right\} + e^{-i(a-b)t} \left\{ |I\rangle - |II\rangle \right\} \right] \\ &= e^{-iat} \left[\cos(bt) |I\rangle - i \sin(bt) |II\rangle \right] . \end{aligned} \tag{271}$$

Eq. (271) has all the interesting stuff of two-level quantum mechanics. It tells you that at a later time, the state is no longer a pure $|I\rangle$; it is an admixture of the two states. Note the oscillatory cosine and sine functions; that is why this is called as oscillation. The probability of finding the state at $|I\rangle$ is maximum at $bt = 0, n\pi$: it is entirely $|I\rangle$, and there is no $|II\rangle$ component. The opposite happens for $bt = (n + \frac{1}{2})\pi$. The total probability is $\cos^2(bt) + \sin^2(bt) = 1$, which it should be. Also note that the oscillation frequency depends on the transition probability b; the higher this probability is, the faster the oscillation.

Now let us take a two-level system where both states can decay to some other final state. A good example is the pair of states K^0 and $\overline{K}^0$: both of them can decay to $\pi^+\pi^-$. Strictly spaking, the Hilbert space is no longer two-dimensional, as one should take into account these final states too, but we can continue the oscillation studies nevertheless. If the state decays, the probability is no longer conserved, and the Hamiltonian need not be hermitian anymore. Thus, both a and b can be complex.

Suppose we write

$$a = a_r - i a_i , \quad b = b_r - i b_i , \tag{272}$$

where the imaginary parts are responsible for decay (we will discuss later how). The lifetimes of the two states, which are proportional to $(a_i + b_i)^{-1}$ and $(a_i - b_i)^{-1}$, are different too. Let us consider a simplifying case: $a = b$. This renders one of the eigenvalues to be zero, so that state, $|2\rangle$, neither evolves nor decays: $|2(t)\rangle = |2(0)\rangle$. One can write (271) as

$$\psi(t) = \frac{1}{2}\left[\left\{e^{-2a_i t}e^{-i2a_r t} + 1\right\}|I\rangle + \left\{e^{-2a_i t}e^{-i2a_r t} - 1\right\}|II\rangle\right] . \tag{273}$$

Thus, the probability of getting the state as $|I\rangle$ and $|II\rangle$ are, respectively,

$$\begin{aligned}|\langle I|\psi\rangle|^2 &= \frac{1}{4}\left[1 + e^{-4a_i t} + 2e^{-2a_i t}\cos(2a_r t)\right] ,\\ |\langle II|\psi\rangle|^2 &= \frac{1}{4}\left[1 + e^{-4a_i t} - 2e^{-2a_i t}\cos(2a_r t)\right] .\end{aligned} \tag{274}$$

Note that apart from the oscillation, there is an exponentially falling envelope. Also, as $t \to \infty$, both probabilities tend to $\frac{1}{4}$. This is easy to understand: half of the particles, being in $|1\rangle$, decayed down, while the other half, in $|2\rangle$, remained as they were. After a long time, when the oscillation effects become unimportant, the probabilities of being in $|I\rangle$ and $|II\rangle$ are equally distributed, and hence $\frac{1}{4}$.

Q. Get an analogous expression of (274) when $a \neq b$, in particular, $a_r > b_r$, $a_i > b_i$. Show that as $t \to \infty$, both probabilities go to zero.

7.6 CP violation in the K system

In 1964, Christenson, Cronin, Fitch, and Turlay observed something rather surprising. They prepared a beam of neutral K mesons. This contains an admixture of both K_L and K_S. While the lifetime of K_S is 0.895×10^{-10} s, that of K_L is almost 600 times as large, *viz.*, 5.12×10^{-8} s, so the K_S component decays quickly and what remains is a beam of pure K_L.

Note that K^0 and $\overline{K}^0$ are not mass eigenstates. There is a nontrivial one-loop diagram, often called the *box* diagram because it resembles a box, that mixes K^0 and $\overline{K}^0$, and hence all our deductions for the two-level system will hold here. We will discuss and evaluate the box diagram for the B mesons later; it suffices to say that whereas K^0 and $\overline{K}^0$ can oscillate into each other, the mass eigenstates K_L and K_S, by definition, cannot.

The collaboration observed that a small number of K_L mesons decay to 2π instead of 3π. The number is small, about two parts in a thousand, but it is definitely there. This, then, implies that not only parity but also the combined symmetry CP is violated in the weak interaction, though the violation of the latter is small and by no means maximal like the parity violation.

If CP is violated, the CP eigenstates K_1 and K_2 can no longer be considered as mass eigenstates. Rather, we will define

$$K_L = \frac{K_1 + \bar\epsilon K_2}{\sqrt{1+|\bar\epsilon|^2}} , \quad K_S = \frac{K_2 + \bar\epsilon K_1}{\sqrt{1+|\bar\epsilon|^2}} , \tag{275}$$

where $\bar\epsilon$ is a small number, and the denominator just normalises the states to unit norm. However, $\bar\epsilon$ depends on the CP convention that we take, and such convention-dependent quantities cannot be a measure of physical observables. However, the quantity ϵ_K, defined as

$$\epsilon_K = \frac{A(K_L \to 2\pi(I=0))}{A(K_S \to 2\pi(I=0))} \tag{276}$$

is measurable. This can be related with $\bar\epsilon$, and the measured value is $(2.280 \pm 0.013) \times 10^{-3}$. The $I = 0$ part means that we take the amplitude when the final state pions are at $I = 0$ (remember that there can be $I = 1$ and $I = 2$ combinations too).

Thus, weak interaction violates CP. The value of ϵ_K tells us that the violation is small, at least in the K system. The CPT theorem, that for any local axiomatic field theory the combined operation CPT is a good symmetry, tells us that time reversal is also violated by the same amount, so weak interaction can in principle discriminate between past and future. Two questions arise.

First, is it good or bad to have CP violation? We know that the universe has much more protons than antiprotons; it is heavily lopsided in favour of matter — one baryon for about 10^{10} photons. How do we know that? If there were antiproton stars and galaxies, we would have observed huge γ-ray bursts when they collide and annihilate with a matter galaxy, but we have not seen anything like that so far. It was shown by Sakharov that to generate such a baryon asymmetry in the universe, it is a must that the early universe had CP violation built in (there are other conditions too); so without CP violation, we would never have been here.

The second question, which we are going to answer now, is how to incorporate CP violation in the weak interaction, now that we know that the Lorentz structure is $V - A$ and there is no scope to violate CP there.

The only way out is to make the accompanying coupling constant complex. Suppose the gauge coupling g is complex. In that case the hermitian conjugate will contain g^*. However, CP conjugation does not affect the coupling, so the CP-conjugated Lagrangian will be different from the original one, and we can have a source of CP-violation.

However, the gauge couplings cannot be complex — that will spoil the hermitian nature of the covariant derivatives. But we have the quark mixing elements too; for two generations they are parametrised with $\cos\theta$ and $\sin\theta$. Thus there cannot be any CP violation if there were only two generations of quark. Fortunately, there are three, and that opens a window to CP violation, which was first identified by Kobayashi and Maskawa.

Q. Show explicitly that a complex coupling constant is essential for CP violation in a $V - A$ theory.

7.7 The Cabibbo-Kobayashi-Maskawa matrix

Let us rcall how the quark masses appeared in the SM. For a single generation, the Yukawa term, say for the down quark, can be written as $-h_d\bar{Q}_L d_R\Phi$ where Q is the quark doublet and d is the singlet, and Φ is the SM Higgs doublet. After shifting the vacuum, we get the mass term as $-m_d\bar{d_L}d_R + h.c.$ where $m_d = h_d v/\sqrt{2}$, v being the VEV of the Higgs field.

For three generations, we change our notations slightly. The flavour or weak basis is not the physical basis, so the mass matrix and hence the Yukawa coupling matrix need not be diagonal. With the primes for the weak basis fields, we can write the down-type Yukawa matrix as

$$\mathcal{L}^d_Y = -\tilde{h}^d_{ij}\bar{Q'}_{Li}d'_{Rj}\Phi + h.c. \tag{277}$$

where i and j runs from 1 to 3, over the generations. The (ij)-th element of the down-type mass matrix is thus

$$m^d_{ij} = \tilde{h}^d_{ij}v/\sqrt{2}\,. \tag{278}$$

Similarly, we can generate the mass matrix for the up-type quarks with $\tilde{\Phi} = i\sigma_2\Phi^*$ and write

$$m^u_{ij} = \tilde{h}^u_{ij}v/\sqrt{2}\,. \tag{279}$$

Note that the matrices $\tilde{h}^d$ and $\tilde{h}^u$ can be completely different.

Let us now rotate to the mass basis, defined as

$$\begin{aligned} d'_R &= \mathcal{D}_R d_R \Longrightarrow d'_{Rj} = (\mathcal{D}_R)_{jl}\, d_{Rl}\,, \\ d'_L &= \mathcal{D}_L d_L \Longrightarrow d'_{Li} = (\mathcal{D}_L)_{ip}\, d_{Lp}\,, \end{aligned} \tag{280}$$

where the matrices $\mathcal{D}_R$ and $\mathcal{D}_L$ rotate the right- and left-chiral down quark fields respectively, and again note that they need not be equal. The last equation implies $\bar{d}'_{Li} = \bar{d}_{Lp}(\mathcal{D}^\dagger_L)_{pi}$, and so the mass term can be rewritten as

$$\mathcal{L}^d_{mass} = -\frac{v}{\sqrt{2}}\bar{d}_{Lp}(\mathcal{D}^\dagger_L)_{pi}\tilde{h}^d_{ij}(\mathcal{D}_R)_{jl}d_{Rl} + h.c. \tag{281}$$

One can always choose the unitary matrices $\mathcal{D}_R$ and $\mathcal{D}_L$ in such a way that they diagonalise $\tilde{h}$; in fact, if they rotate the weak basis to the mass basis, the mass matrix must be diagonalised by definition. Such a transformation is known as a *biunitary transformation.* One can perform an identical exercise for the up-type quarks with the unitary matrices $\mathcal{U}_L$ and $\mathcal{U}_R$.

We thus define the diagonal matrices

$$h^d = \mathcal{D}^\dagger_L \tilde{h}^d \mathcal{D}_R\,, \quad h^u = \mathcal{U}^\dagger_L \tilde{h}^u \mathcal{U}_R\,, \tag{282}$$

with the quark masses $m_i = h_{ij}\delta_{ij}v/\sqrt{2}$.

One can write the CC Lagrangian as

$$\begin{aligned} \mathcal{L}_{cc} &= -\frac{g}{\sqrt{2}}\bar{u}_{Li}(\mathcal{U}_L)^\dagger_{ip}(\mathcal{D}_L)_{pj}\gamma^\mu P_L d_j W^+_\mu + h.c. \\ &= -\frac{g}{\sqrt{2}}V_{ij}\bar{u}_{Li}\gamma^\mu P_L d_j W^+_\mu + h.c. \end{aligned} \tag{283}$$

where $V = \mathcal{U}^\dagger_L \mathcal{D}_L$ is a unitary matrix and controls the quark mixing. This is known as the *Cabibbo-Kobayashi-Maskawa matrix* or the CKM matrix. The entire physics of quark flavours is the physics of the CKM matrix.

Note a few points:

- All the $\mathcal{U}$ and $\mathcal{D}$ matrices are unitary, and thus the combinations $\mathcal{D}^\dagger_L\mathcal{D}_L$ or $\mathcal{U}^\dagger_L\mathcal{U}_L$ occuring in NC interactions are unit matrices. This ensures the GIM mechanism: there is no FCNC at the tree-level of the SM.

- There is absolutely no way to determine the individual $\mathcal{U}$ or $\mathcal{D}$ matrices. What one can only determine is the elements of the combination V. Thus, without loss of any generality, one can take $V = \mathcal{D}_L$ and $\mathcal{U}_L = \mathbf{1}$; the rotation is confined in the down-type quark sector, whereas for the up-type quarks, the gauge basis and the stationary basis are assumed to be identical. We could have done the opposite: $\mathcal{U}_L = V^\dagger$ and $\mathcal{D}_L = \mathbf{1}$, or even some intermediate combination. The standard convention is just a reminiscence of the Cabibbo mechanism. It is only the misalignment between the up-type basis and the down-type basis that is important. Also, we cannot have any information about $\mathcal{U}_R$ or $\mathcal{D}_R$.

- For two generations, the CKM matrix reduces to the Cabibbo matrix parametrised by θ. For three generations, there are three real elements and one complex phase in the CKM matrix, and hence the CC Lagrangian leads to CP violation.

7.8 The Wolfenstein parametrisation

The CKM matrix V can be written explicitly as

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{284}$$

where the element V_{ij} gives the strength of the transition from the i-type charge $2/3$ quark to the j type charge $-1/3$ quark. The elements are not arbitrary; all of them can be parametrised by at most four independent parameters, three real angles and a complex phase.

Neutron β-decay measures V_{ud}; the decay of the D mesons, containing a charm quark, to K mesons measures V_{cs}, and both turn out to be close to unity. This is not surprising; the upper left 2×2 block involving only the first two generations is the Cabibbo matrix, and we know $\cos\theta \approx 1$. The top quark is so heavy that it decays before it has time to pick up an antiquark and form a bound state, so there are no top mesons; but all the top quarks that were produced in the laboratory (which means, so far, the Tevatron experiment at Fermilab) decay to bottom quarks, so V_{tb} must also be very close to unity. Thus, V is *almost* a unit matrix, and the probability of generational crossover transitions is relatively low.

V_{us} and V_{cd} are both small; their magnitudes are given, in the first approximation, by the sine of the Cabibbo angle $\sin\theta \approx 0.22$. Just like the s quark, the b quark can only decay through generation-changing charged current, as $m_b \ll m_t$. The dominant decay channel is $b \to c$, which measures V_{cb}. This turns out to be even smaller than V_{us} or V_{cd}, at the same order of (but not equal to) $\sin^2\theta \approx 0.04$. This tells us why the bottom quark, in spite of its mass of about 4.5 GeV, is relatively long-lived; the decay width Γ is multiplied by $|V_{cb}|^2$, and the lifetime is of the order of 10^{-12} seconds, which is easily measurable in experiments. (Show that the lifetime $\tau_b = 1/\Gamma_b$ is indeed of the order of a picosecond.)

The b quark can also decay to an up quark, with smaller probability, because in this case the generation changes by two. This gives us V_{ub}, which is even smaller, of the order of $\sin^3\theta \approx 8\times 10^{-3}$. The constraint that V is unitary helps us to find the other two elements: V_{td} turns out to be of the same order as V_{ub} and $|V_{ts}| \approx |V_{cb}|$. Thus, one can parametrise the CKM matrix as

$$V = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(1-\rho-i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{285}$$

Here $\lambda \approx \sin\theta$ is a smallness parameter by which we parametrise the relative magnitudes of the CKM elements, and the other real constants A, ρ, and η are all $\mathcal{O}(1)$. Note that the phase is confined to V_{ub} and V_{td}. This is purely a convention, but a good point is that this keeps the 2×2 Cabibbo matrix real. This matrix is unitary up to order λ^3 (check it!).

From $VV^\dagger = V^\dagger V = \mathbf{1}$, one can write several constraint equations, like

$$\begin{aligned} V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* &= 0\,, \\ V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* &= 0\,, \\ V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* &= 0\,, \\ V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* &= 0\,, \\ V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* &= 0\,, \\ V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* &= 0\,. \end{aligned} \tag{286}$$

The elements, at least some of them, are complex, so such equations represent triangles in the complex plane. These triangles are known as *unitarity triangles*. However, note that not all of them are readily identifiable as triangles. For example, consider the first and the fourth equations. Each of these two triangles has two sides which are of the order of λ and one which is of the order of λ^5, so they look like very squashed triangles [20]. Similarly, the third and the sixth relationships give two more triangles, each of which has two sides of the order of λ^2 and one of the order of λ^4, so they are squashed triangles too.

If $\eta = 0$, the CKM matrix is real, so the charged current couplings are all real, and there is no CP violation. In that limit, obviously, all triangles collapse to straight lines. Thus, the fact that these triangles are actual triangles with nonzero areas is crucially important for the observation of CP violation. To measure the 'triangleness', one has to measure the angles (to check that they are not zero or π), but that is almost impossible for such a squashed triangle.

However, there are two triangles, relationships 2 and 5, that look like triangles. The reason is that all the sides are of equal order, λ^3. In particular, the second relationship can be measured quite precisely, and that is why the triangle given by that is commonly called *the unitarity triangle.*

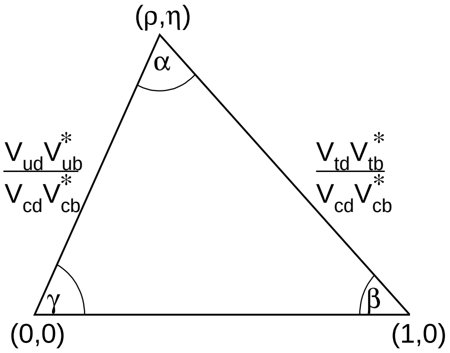
(ρ,η)
α
$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}$
$\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}$
γ
β
(0,0)
(1,0)

Figure 4: *The unitarity triangle.*

The unitarity triangle is shown in figure 7.8. The sides have been scaled down by $|V_{cd}V_{cb}^*| = A\lambda^3$, so that the vertex is at (ρ, η). With the Wolfenstein convention where only V_{td} and V_{ub} are complex, we can write

$$V_{td} = |V_{td}| \exp(-i\beta) , \quad V_{ub} = |V_{ub}| \exp(-i\gamma) . \tag{287}$$

Q. Show that eq. (285) is unitary if we neglect terms of the order of λ^4 or higher.
Q. Convince yourself that the vertex of figure 7.8 is at (ρ, η). Also convince yourself that the phase of V_{td} and V_{ub} are $-\beta$ and $-\gamma$ respectively.
Q. Show that the area of any unitarity triangle is given by

$$\text{Area} = \frac{1}{2}\text{Im}(V_{ip}V_{jq}^*V_{iq}^*V_{jp}) , \tag{288}$$

where i, j are charge 2/3 quarks and p, q are charge $-1/3$ quarks. [*Hint: rotate the triangle so that $V_{ip}V_{iq}^*$ lies along the real axis.*]
Q. Show that all unitarity triangles, regardless of their shapes, have same area. [*Hint: Start from the first relationship of eq. (286) and write the area as $\frac{1}{2}\text{Im}(V_{ud}V_{us}^*V_{cd}^*V_{cs})$. Replace $V_{ud}V_{cd}^*$ using*

---

[20] If we stick to the leading order, eq. (285), the last term is zero, as we neglect terms higher than λ^3, so the triangles become straight lines. However, at higher orders, they are triangles.

the fourth relationship, use the fact that $|V_{us}|^2|V_{cs}|^2$ *is real, and get the area of the third triangle. Similarly, you can move to other triangles.*]
Thus, the quantity $|\mathrm{Im}(V_{ip}V_{jq}^*V_{iq}^*V_{jp})|$ is an invariant for three-generation CKM matrix. This is known as the *Jarlskog invariant* and is a measure of CP violation in the Standard Model.

7.9 $B^0 - \overline{B}^0$ mixing: box diagram

There are four possible diagrams for $B^0 - \overline{B}^0$ mixing in the 't Hooft-Feynman gauge. In the unitary gauge, only the WW box exists.

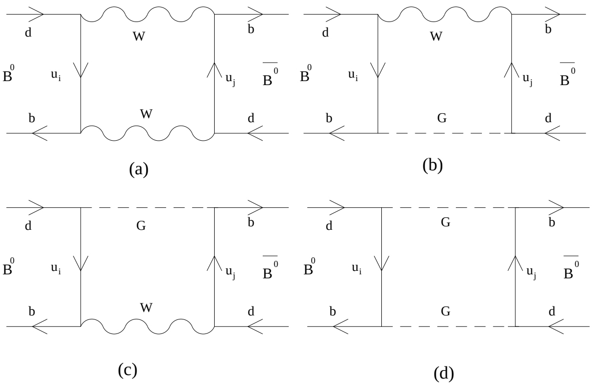

Figure 5: The relevant diagrams for $B^0 - \overline{B}^0$ box in the 't Hooft-Feynman gauge. Only (a) exists in the unitary gauge.

We set the following shorthands:

$$\lambda_i = V_{id}V_{ib}^*\,, \quad D = (k^2 - m_i^2)(k^2 - m_j^2)(k^2 - m_W^2)^2\,, \quad D_E = (k_E^2 + m_i^2)(k_E^2 + m_j^2)(k_E^2 + m_W^2)^2\,, \tag{289}$$

where m_i denotes the mass of the u_i quark ($i = u, c, t$). The subscript E denotes the Euclidean momenta. We also define

$$O_1 = (\bar{b}\gamma^\mu(1-\gamma_5)d)(\bar{b}\gamma_\mu(1-\gamma_5)d) \tag{290}$$

in a singlet-singlet combination of $SU(3)_c$.

For terms like $[\bar{b}\gamma^\mu\gamma^\alpha\gamma^\nu d][\bar{b}\gamma_\nu\gamma_\alpha\gamma_\mu d]$, we use the so-called Greek method: insert another γ matrix in place of $d\bar{b}$ in the middle, perform the usual tricks of γ matrices, and ultimately replace the γ matrix by $d\bar{b}$.

Also note that terms like $k^\mu k^\nu$, in a symmetric integration, can be replaced by $\eta^{\mu\nu}k^2/4$.

Unitary Gauge

The amplitude can be written as

$$\begin{aligned} i\mathcal{M} = \int \frac{d^4k}{(2\pi)^4} \quad & \left[\bar{b}\left(\frac{-ig}{\sqrt{2}}\right)\gamma^\alpha P_L \frac{i}{\not{k} - m_i}\left(\frac{-ig}{\sqrt{2}}\right)\gamma^\mu P_L d\right] \times \\ & \left[\bar{b}\left(\frac{-ig}{\sqrt{2}}\right)\gamma^\nu P_L \frac{i}{\not{k} - m_j}\left(\frac{-ig}{\sqrt{2}}\right)\gamma^\beta P_L d\right] \times \\ & \lambda_i\lambda_j\left(\frac{-i}{k^2 - m_W^2}\right)^2\left(\eta_{\mu\nu} - \frac{k_\mu k_\nu}{m_W^2}\right)\left(\eta_{\alpha\beta} - \frac{k_\alpha k_\beta}{m_W^2}\right) \end{aligned}$$

$$
\begin{aligned}
= \quad & \frac{g^4}{4}\lambda_i\lambda_j \int \frac{d^4k}{(2\pi)^4}\frac{1}{D}\left[\bar{b}\gamma^\alpha \not{k}\gamma^\mu P_L d\right]\left[\bar{b}\gamma^\nu \not{k}\gamma^\beta P_L d\right] \times \\
& \left(\eta_{\mu\nu}\eta_{\alpha\beta} - \eta_{\mu\nu}\frac{k_\alpha k_\beta}{m_W^2} - \eta_{\alpha\beta}\frac{k_\mu k_\nu}{m_W^2} + \frac{k_\mu k_\nu k_\alpha k_\beta}{m_W^4}\right) \qquad (291)
\end{aligned}
$$

The different terms, after some not-too-difficult manipulation, come out as

$$
\begin{aligned}
\text{Term1} &= \left[\bar{b}\gamma^\alpha \not{k}\gamma^\mu P_L d\right]\left[\bar{b}\gamma^\nu \not{k}\gamma^\beta P_L d\right]\eta_{\mu\nu}\eta_{\alpha\beta} \\
&= \left[\bar{b}\gamma^\mu P_L d\right]\left[\bar{b}\gamma_\mu P_L d\right].k^2, \qquad (292) \\
\text{Term2} &= -\eta_{\mu\nu}\frac{k_\alpha k_\beta}{m_W^2}\left[\bar{b}\gamma^\alpha \not{k}\gamma^\mu P_L d\right]\left[\bar{b}\gamma^\nu \not{k}\gamma^\beta P_L d\right] \\
&= -\frac{k^4}{m_W^2}\left[\bar{b}\gamma^\mu P_L d\right]\left[\bar{b}\gamma_\mu P_L d\right], \qquad (293) \\
\text{Term3} &= \text{Term2}, \qquad (294) \\
\text{Term4} &= \frac{1}{m_W^4}\left[\bar{b}\not{k}\not{k}\not{k}P_L d\right]\left[\bar{b}\not{k}\not{k}\not{k}P_L d\right] \\
&= \frac{k^6}{4m_W^4}\left[\bar{b}\gamma^\mu P_L d\right]\left[\bar{b}\gamma_\mu P_L d\right]. \qquad (295)
\end{aligned}
$$

We have used $\not{k}\not{k} = k^2$ to reduce the number of γ-matrices in the last three equations.

Thus,

$$
\begin{aligned}
i\mathcal{M} &= \sum_i\sum_j \frac{g^4}{4(2\pi)^4}\lambda_i\lambda_j \int \frac{d^4k}{D}\left(k^2 - \frac{2k^4}{m_W^2} + \frac{k^6}{4m_W^4}\right) \times \frac{1}{4}O_1 \\
&= \frac{g^4}{16(2\pi)^4}\sum_i\sum_j \lambda_i\lambda_j O_1 \int \frac{d^4k}{D}\left(k^2 - \frac{2k^4}{m_W^2} + \frac{k^6}{4m_W^4}\right). \qquad (296)
\end{aligned}
$$

The integrals can be easily performed by a rotation to the Euclidean space, with $k^2 \to -k_E^2$, $d^4k \to id^4k_E = i\pi^2 k_E^2 dk_E^2$.

$$
I_1 = \int d^4k \frac{k^2}{D} =_{M\to E} -i\pi^2 \int \frac{k^4 dk^2}{D_E} \qquad (297)
$$

where we have written k for k_E in the last step. The integrand can be written as

$$
\frac{k^4}{D_E} = \frac{1}{(k^2+m_W^2)^2} - \frac{m_i^2}{(k^2+m_i^2)(k^2+m_W^2)^2} - \frac{m_j^2}{(k^2+m_j^2)(k^2+m_W^2)^2} + \frac{m_i^2 m_j^2}{D_E}. \qquad (298)
$$

The first three terms vanish due to unitarity of the CKM matrix: either $\sum_i \lambda_i = 0$ or $\sum_j \lambda_j = 0$. Thus, we have

$$
\begin{aligned}
I_1 &= -i\pi^2 m_i^2 m_j^2 \int \frac{dk^2}{(k^2+m_i^2)(k^2+m_j^2)(k^2+m_W^2)^2} \equiv -i\pi^2 m_i^2 m_j^2 \tilde{I}_1, \\
I_2 &= -\frac{2}{m_W^2}\int d^4k \frac{k^4}{D} =_{M\to E} -\frac{2}{m_W^2} i\pi^2 \int \frac{k^6 dk^2}{D_E} = -\frac{2}{m_W^2} i\pi^2 m_i^2 m_j^2 \int \frac{k^2 dk^2}{D_E}
\end{aligned}
$$

$$\begin{aligned} &\equiv -\frac{2}{m_W^2} i\pi^2 m_i^2 m_j^2 \tilde{I}_2, \\ I_3 &= \frac{1}{4m_W^4}\int d^4k \frac{k^6}{D} =_{M\to E} -i\pi^2 \frac{1}{4m_W^4}\int \frac{k^8 dk^2}{D_E} = -i\pi^2 \frac{1}{4m_W^4} m_i^2 m_j^2 \int \frac{k^4 dk^2}{D_E} \\ &\equiv -i\pi^2 \frac{1}{4m_W^4} m_i^2 m_j^2 \tilde{I}_3. \end{aligned} \tag{299}$$

Therefore

$$i\mathcal{M} = \sum_i \sum_j \frac{g^4}{16(2\pi)^4}\lambda_i\lambda_j O_1(-i\pi^2) m_i^2 m_j^2 \left[\tilde{I}_1 + \frac{2}{m_W^2}\tilde{I}_2 + \frac{1}{4m_W^4}\tilde{I}_3\right]. \tag{300}$$

For $i = j = t$ (say), we have

$$\tilde{I}_1 + \frac{2}{m_W^2}\tilde{I}_2 + \frac{1}{4m_W^4}\tilde{I}_3 = \frac{1}{m_W^6 x_t^2} S_0(x_t) \tag{301}$$

where $x_t = m_t^2/m_W^2$, and

$$S_0(x) = \frac{4x - 11x^2 + x^3}{4(1-x)^2} - \frac{3x^3 \log x}{2(1-x)^3}. \tag{302}$$

Thus,

$$\begin{aligned} i\mathcal{M} &= \frac{g^4}{16(2\pi)^4}\lambda_t^2 O_1(-i\pi^2) m_t^4 \frac{1}{m_W^6 x_t^2} S_0(x_t) \\ &= -i\frac{G_F^2}{8\pi^2}\lambda_t^2 m_W^2 S_0(x_t) O_1. \end{aligned} \tag{303}$$

For $B^0 - \overline{B}^0$ mixing, only the top-mediated loop is numerically important. However, for completeness, let us just quote the result when two different quarks flow inside the box. Suppose $i = t$ and $j = c$. In this case

$$\tilde{I}_1 + \frac{2}{m_W^2}\tilde{I}_2 + \frac{1}{4m_W^4}\tilde{I}_3 = \frac{N_1}{D_1}, \tag{304}$$

where

$$\begin{aligned} N_1 = \quad & \Big[(m_c^2 - m_W^2)^2 (m_t^4 - 8m_t^2 m_W^2 + 4m_W^4)\log m_t^2 - (c \leftrightarrow t) - 3(m_t^2 - m_c^2) m_W^2 \times \\ & \Big\{(m_t^2 - m_W^2)(m_c^2 - m_W^2) + \Big(2m_c^2 m_t^2 - m_c^2 m_W^2 - m_t^2 m_W^2\Big)\log m_W^2\Big\}\Big], \end{aligned} \tag{305}$$

and

$$D_1 = 4m_W^4 (m_t^2 - m_c^2)(m_t^2 - m_W^2)^2 (m_c^2 - m_W^2)^2. \tag{306}$$

After some trivial algebraic manipulations, the amplitude comes out to be

$$\begin{aligned} i\mathcal{M} &= \frac{g^4}{16(2\pi)^4}\lambda_t\lambda_c O_1(-i\pi^2) m_t^2 m_c^2 (N_1/D_1) \\ &= -i\frac{G_F^2}{8\pi^2}\lambda_t\lambda_c m_W^2 S_0(x_t, x_c) O_1, \end{aligned} \tag{307}$$

where $S_0(x_t, x_c)$ is given by

$$S_0(x_t, x_c) = x_c\left[\log\frac{x_t}{x_c} - \frac{3x_t}{4(1-x_t)} - \frac{3x_t^2 \log x_t}{4(1-x_t)^2}\right], \tag{308}$$

and we have neglected all higher powers of x_c.

This overall amplitude is to be multiplied by 2 to take into account the crossed diagrams, where the quarks flow horizontally and bosons vertically. This is true even for the $t-c$ box. However, unless explicitly stated otherwise, we will always neglect everything but the $t-t$ box for the SM part. (But note that they play an extremely crucial role for GIM cancellation!)

This generates the Wilson coefficient, C_1, at the scale m_W. Now, if the Hamiltonian is of the form

$$\mathcal{H} = C_1 O_1 \tag{309}$$

Then the amplitude is given by

$$i\mathcal{M} = -i(4C_1)O_1 \tag{310}$$

(actually, the expectation value of O_1 between B^0 and $\bar{B^0}$). The negative sign comes from the fact that the amplitude comes from the effective Lagrangian, which is the negative of the effective Hamiltonian. The factor i comes from $S = \exp(i\int \mathcal{L}d^4x)$. There are two ways of annihilating a d quark, and two ways of creating a b quark, so that gives an extra factor of $2\times 2 = 4$. Thus, the Hamiltonian is

$$\mathcal{H}^{\Delta B=2} = \frac{G_F^2}{16\pi^2}\lambda_t^2 m_W^2 S_0(x_t) O_1. \tag{311}$$

This has to be corrected for short-distance QCD.

’t Hooft - Feynman Gauge

In this gauge all four diagrams (a)-(d) are present, but the W propagator is modified — there is no $k_\mu k_\nu$ term, and correspondingly, only I_1 is present.

To evaluate diagrams (b)-(d), we need the coupling of the Goldstone to quarks, which is given by the Lagrangian (see, *e.g.*, Cheng and Li)

$$\mathcal{L}_G = -\frac{g}{2\sqrt{2}m_W}\left[V_{ij}\bar{u}_i\Big\{m_{d_j}(1+\gamma_5) - m_{u_i}(1-\gamma_5)\Big\}d_j + V_{ij}^*\bar{d}_j\Big\{m_{d_j}(1-\gamma_5) - m_{u_i}(1+\gamma_5)\Big\}u_i\right]. \tag{312}$$

We neglect all quarks masses except that of the top quark. This gives for the diagram (d):

$$i\mathcal{M}_{GG} = -\frac{i}{32\pi^2}G_F^2\lambda_t^2 m_t^4 O_1 \int \frac{k^4 dk^2}{(k^2+m_t^2)^2(k^2+m_W^2)^2} \tag{313}$$

(remember our definition of O_1). This is to be multiplied by 2 for another possible Wick contraction.

Diagrams (b) and (c) contribute equally, so it is enough to calculate one and multiply by 4 (2 times 2 for another Wick contraction). Neglecting all quark masses except m_t gives

$$i\mathcal{M}_{WG} = -\frac{i}{8\pi^2}G_F^2\lambda_t^2 m_W^2 m_t^4 O_1 \int \frac{k^2 dk^2}{(k^2+m_t^2)^2(k^2+m_W^2)^2}. \tag{314}$$

Thus, the total contribution is

$$\begin{aligned}\mathcal{M} &= 2\mathcal{M}_{WW} + 4\mathcal{M}_{WG} + 2\mathcal{M}_{GG} \\ &= -\frac{i}{\pi^2}G_F^2\lambda_t^2 m_t^4 O_1\left[\frac{1}{16}\tilde{I}_3^{\,i=j=t} + \frac{1}{2}m_W^2\tilde{I}_2^{\,i=j=t} + \frac{1}{4}m_W^4\tilde{I}_1^{\,i=j=t}\right].\end{aligned} \tag{315}$$

This gives us the same effective Hamiltonian, which it must, when one follows the prescription of multiplying the amplitude by $-(1/4)$ to get $\mathcal{H}_{eff}$.

7.10 CP violation in the B system

A complex coupling constant is not enough to observe CP violation. We observe CP violation if the rates of some initial state to some final state $i \to f$, and its CP-conjugate process $\bar{i} \to \bar{f}$, are not equal. In other words, we can take, as a measure of CP asymmetry, the ratio

$$A_{CP} = \frac{\Gamma(i \to f) - \Gamma(\bar{i} \to \bar{f})}{\Gamma(i \to f) + \Gamma(\bar{i} \to \bar{f})}\,. \tag{316}$$

All other factors being equal, this means that the amplitudes squared must be different.

Consider the decay $B^+ \to f$ and $B^- \to \bar{f}$, where f is some mesonic final state (say $J/\psi K^+$, where J/ψ is the lowest-lying spin-1 bound state of $c\bar{c}$ [21]). If the final state consists of hadrons, there might be a further scattering by the exchange of one or more gluons; this is known as *final-state* interaction. Just like any other scattering, this introduces a phase shift between the initial and final states and is called the *strong phase* of the amplitude. Under CP conjugation, the weak phase — the one associated with the CKM matrix — changes sign but the strong phase doesn't, as gluons couple equally to quarks and antiquarks. Thus, we can write the amplitude as

$$M_{B^+\to f} = a\exp(i\alpha)\exp(i\delta)\,, \quad \bar{M}_{B^-\to\bar{f}} = a\exp(-i\alpha)\exp(i\delta)\,, \tag{317}$$

where a is a real number and α and δ are the weak and strong phases respectively. Obviously, $|M|^2 = |\bar{M}|^2$, so there is no CP violation.

Now suppose that there are two different amplitudes for each decay. They can involve different intermediate quark states and hence different weak and strong phases. Let us write

$$\begin{aligned} M_{B^+\to f} &= a_1\exp(i\alpha_1)\exp(i\delta_1) + a_2\exp(i\alpha_2)\exp(i\delta_2)\,,\\ \bar{M}_{B^-\to\bar{f}} &= a_1\exp(-i\alpha_1)\exp(i\delta_1) + a_2\exp(-i\alpha_2)\exp(i\delta_2)\,, \end{aligned} \tag{318}$$

so that

$$\begin{aligned} |M|^2 &= a_1^2 + a_2^2 + 2a_1a_2\cos(\Delta\alpha + \Delta\delta)\,,\\ |\bar{M}|^2 &= a_1^2 + a_2^2 + 2a_1a_2\cos(-\Delta\alpha + \Delta\delta)\,, \end{aligned} \tag{319}$$

where $\Delta\alpha = \alpha_2 - \alpha_1$, $\Delta\delta = \delta_2 - \delta_1$. Therefore

$$A_{CP} \propto -4a_1a_2\sin(\Delta\alpha)\sin(\Delta\delta)\,. \tag{320}$$

So, in order to observe CP violation, we must have *at least two amplitudes with different weak phases and different strong phases.*

If the final state $|f\rangle$ is such that it can be accessed from both B^0 and $\overline{B}^0$, the second amplitude is automatically provided by mixing: $B^0 \to f$ and $B^0 \to \overline{B}^0 \to f$. This is a cleaner situation as the extracted observables do not depend upon the strong phase difference, which is a priori not calculable. The angles of the UT, in particular β, were extracted from such decays.

Q. Check that the following states can be accessed from both B^0 and $\overline{B}^0$: $\pi^+\pi^-$, $J/\psi K_S$, D^+K^- (K_S is a linear superposition of K^0 and $\overline{K}^0$, $D^+ \equiv c\bar{d}$). Note that the final state need not be a CP

---

[21] This is the only particle that is denoted by two letters. The reason is interesting. It was simultaneously discovered by two different groups; Sam Ting, the leader of one group, called it J, Burton Richter, the leader of another, callled it ψ. The name is to honour both; they also shared a Nobel prize.

eigenstate.
Q.(*) Show that the phase of the $B^0 - \overline{B}^0$ box amplitude, with top quark running inside, is $\phi_M = -2\beta$. For a final state $|f\rangle$ which is a CP eigenstate, the measured angle is the argument of

$$\exp(i\phi_M)\frac{A(\overline{B}^0 \to f)}{A(B^0 \to f)}\,. \tag{321}$$

Show that if $f = \pi^+\pi^-$, the angle measured is α while if it is $J/\psi K_S$, the angle measured is β. While you have used the Wolfenstein representation, show that the result is representation-independent.

8 Neutrinos

In the Standard Model, there are three neutrinos: ν_e, ν_μ, and ν_τ, associated with the charged leptons e, μ, and τ of the three generations through charged-current interactions. Thus, a charged current interaction can change a neutrino to its corresponding charged lepton and vice versa; there is no generational crossover like the quarks. In other words, this means that the mass and the weak bases are identical for the lepton sector. That is easy to understand: if two neutrinos ν_1 and ν_2 are both massless, any combination $a\nu_1 + b\nu_2$ must also be a stationary state with mass eigenvalue zero; but if all such combinations are stationary states, we can start from the beginning with them, and identify the mass basis with the weak basis. This is why, if the neutrinos are massless (or all have the same mass), the individual lepton numbers are conserved.

For massless neutrinos, we have already derived the following:

- It is possible to write the wavefunctions as 2-component Weyl spinors. The Hamiltonian $H = \gamma^0\gamma^i p_i$ commutes with γ_5, so chirality is a good quantum number. As the neutrinos are massless, they travel with the velocity of light, and it is impossible to overtake them; thus helicity and chirality are identical.

- $\nu_L = P_L\nu$ and $\nu_R = P_R\nu$, where $P_{L(R)} = \frac{1}{2}(1 - (+)\gamma_5)$, are in two different representations of the Lorentz group: $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$. Under the parity transformation, these two representations change into one another, so ν_L goes to ν_R.

- Neutrinos have only weak interaction; but that violates parity, so ν_L and ν_R need not be on the same footing. In fact, as was shown by Goldhaber and his collaborators, neutrinos are completely left-handed, which established the $V - A$ nature of the weak interaction. Thus, ν_R is indeed a very bizarre object; it does not feel any force except that of gravity.

- For massless neutrinos, we do not need the ν_R to exist at all! The only objects that exist are the two-component spinor ν_L, and its antiparticle $\bar{\nu}_R = \bar{\nu}P_L$. The charged current weak interaction can be written as

$$\mathcal{L}_{cc} = -\frac{g}{2\sqrt{2}}\left(\bar{\nu}\gamma_\mu(1-\gamma_5)eW^{\mu+} + \text{h.c.}\right)\,. \tag{322}$$

8.1 The solar neutrino problem

Neutrinos are produced in the pp chain which fuses four protons to an α particle. There are three ways the pp chain can go: most of the neutrinos are produced in $p + p \to d + e^+ + \nu_e$, a smaller fraction in ${}^7Be + e^- \to {}^7Li + \nu_e$, and an even smaller fraction in ${}^8B \to {}^8Be + e^+ + \nu_e$. The Standard Solar Model predicts the flux of these neutrinos; in units of cm^{-2}s^{-1}, the flux of pp, ${}^7Be$ and ${}^8B$ neutrinos are, approximately, 6×10^{10}, 5×10^9, and 6×10^6 respectively. What we get on earth is something between $\frac{1}{2}$ and $\frac{1}{3}$ of the expected flux. Some of the electron neutrinos must be getting lost somewhere else. This is the solar neutrino problem.

The electron neutrinos are detected in various ways. For example, the gallium or chlorine detectors use

$$ {}^{71}Ga + \nu_e \to {}^{72}Ge + e^- \,, \quad {}^{37}Cl + \nu_e \to {}^{37}Ar + e^- \,. \tag{323} $$

Measuring the quantities of germanium or argon produced, one can estimate the neutrino flux. The water Cerenkov detectors use the scattering $\nu_e + e \to e + \nu_e$, a charged-current process (with a limited sensitivity to neutral current scattering too). So they can detect only ν_e and not ν_μ, muons are too massive to be produced in Ga or Cl detectors, and there are no free muons in water to give a charged current scattering.

Thus, if some ν_e on their way to earth are converted to ν_μ (or something else), they won't be detected, and we will see a deficiency. But that means neutrinos are massive, and ν_e or ν_μ are not stationary states. We have to play with two-level quantum mechanics again.

Consider the non-stationary basis (ν_e, ν_μ) and the stationary basis (ν_1, ν_2). They must be related by

$$ \begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \,, \tag{324} $$

where θ is analogous to the Cabibbo angle. If the states ν_1 and ν_2 have energies E_1 and E_2 respectively, we can write

$$ E_1 = \sqrt{|\mathbf{p}_1|^2 + m_1^2} \,, \quad E_2 = \sqrt{|\mathbf{p}_2|^2 + m_2^2} \,. \tag{325} $$

Neutrinos are almost massless and moving almost at the speed of light, so we can write $|\mathbf{p}_1| = |\mathbf{p}_2| = p$, and

$$ E_1 \approx p + \frac{m_1^2}{2p} = p + \frac{m_1^2}{2E} \,, \quad E_2 \approx p + \frac{m_2^2}{2p} = p + \frac{m_2^2}{2E} \,. \tag{326} $$

Thus, after a time t from when a ν_e was produced in the sun, its wavefunction will be

$$ \nu_e(t) = e^{-ipl} \left[\nu_1 \cos\theta e^{-im_1^2 l/2E} + \nu_2 \sin\theta e^{-im_2^2 l/2E} \right] \,, \tag{327} $$

where we have put $t = l$, the distance travelled, as the neutrinos are moving almost at the velocity of light. Writing ν_1 and ν_2 in terms of ν_e and ν_μ, we get the probability that an electron neutrino will be converted to a muon neutrino:

$$ P_{e\mu}(l) = |\langle \nu_e(t) | \nu_\mu \rangle|^2 = \left| \cos\theta \sin\theta \left(-e^{-im_1^2 l/2E} + e^{-im_2^2 l/2E} \right) \right|^2 = \sin^2 2\theta \, \sin^2 \left(\frac{\Delta m^2 l}{4E} \right) \,, \tag{328} $$

where $\Delta m^2 = m_2^2 - m_1^2$.

Note that this is not the only mechanism leading to $\nu_e \leftrightarrow \nu_\mu$ oscillation. There is another effect mediated by the ambient electrons inside the sun. While (328) does not predict the sign of Δm^2, due to the second effect we know that $\Delta m^2 > 0$, or $m_2^2 > m_1^2$. Also note that once the neutrinos have mass, the individual lepton numbers are no longer conserved, just like strangeness.

Q. pp and ${}^8B$ neutrinos have an energy distribution but those coming from ${}^7Be$ have a sharp peak. Why?

Q. Solar photons show a blackbody spectrum. Why not solar neutrinos?
Q. Show that in more conventional units of Δm^2 (eV^2), l (m), and E (MeV), Eq. (328) becomes

$$P_{e\mu}(l) = \sin^2 2\theta \, \sin^2 \left(\frac{1.3 \Delta m^2 \, l}{E} \right) . \tag{329}$$

8.2 The atmospheric neutrino problem

Cosmic rays produce charged π or K mesons in the upper atmosphere. They undergo weak decay: $\pi^+ (K^+) \to \mu^+ + \nu_\mu$. The electron channels, as we have seen, are suppressed by chirality. Muons decay via $\mu^+ \to e^+ + \nu_e + \bar{\nu_\mu}$. Thus, muon and electron neutrinos should come with a ratio of $2 : 1$ (we include antineutrinos too in this count). There is no scope for confusion with solar neutrinos, the energies are vastly different: at most a few MeV for solar and about a hundred MeV for atmospheric neutrinos.

The number of electron neutrinos that we get are as expected (the length that they get to oscillate into other types of neutrinos is so small that the oscillation probability is negligible). But the muon neutrino numbers fall short by about 50%, so ν_μ must be going to something else, probably ν_τ.

The expression is similar to that of (328), with $\Delta m^2 = m_3^2 - m_2^2$ (so we have to take into account a three-level system, but both these oscillations can approximately be described by two two-level systems). We do not know the sign of Δm^2.

8.3 Neutrino mixing and hierarchy

In the SM with three neutrinos, we can write a mixing matrix as

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} c_{12} c_{13} & s_{12} c_{13} & s_{13} e^{-i\delta} \\ -s_{12} c_{23} - c_{12} s_{23} s_{13} e^{i\delta} & c_{12} c_{23} - s_{12} s_{23} s_{13} e^{i\delta} & s_{23} c_{13} \\ s_{12} s_{23} - c_{12} c_{23} s_{13} e^{i\delta} & -c_{12} s_{23} - s_{12} c_{23} s_{13} e^{i\delta} & c_{23} c_{13} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} , \tag{330}$$

where $s_{ij} \equiv \sin\theta_{ij}$, $c_{ij} \equiv \cos\theta_{ij}$, and δ is a CP-violating phase, just like the quark mixing (note that here also we need only four parameters, three real angles and a phase, to parametrize the matrix). This matrix is known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, something completely analogous to the CKM matrix, but the parametrization is different from the Wolfenstein one. Note that if $\theta_{13} = 0$, there would not have been any leptonic CP violation.

So far, we know the following neutrino mixing parameters (Particle Data Group 2012):

$$\begin{gathered} \Delta m_{21}^2 = (7.50 \pm 0.20) \times 10^{-5}\ \mathrm{eV}^2 , \quad |\Delta m_{32}^2| = (2.32^{+0.12}_{-0.08}) \times 10^{-3}\ \mathrm{eV}^2 , \\ \sin^2 2\theta_{12} = 0.857 \pm 0.024 , \quad \sin^2 2\theta_{23} > 0.95 , \quad \sin^2 2\theta_{13} = 0.098 \pm 0.013 . \end{gathered} \tag{331}$$

That θ_{13} is definitely nonzero was measured very recently by two experiments. While this opens the possibility of leptonic CP violation, so far we have no idea about δ — CP violation is yet to be observed in the lepton sector. The angles are determined only with discrete ambiguities; only $\sin^2 2\theta_{ij}$ has been determined and not θ_{ij}. (How many values between 0 and 2π for an angle θ are possible that will give identical $\sin^2 2\theta$?) Also, sign of Δm_{32}^2 is unknown. This leads to two different hierarchies, both with a small solar splitting and a large atmospheric splitting:

1. Normal hierarchy: $m_1 < m_2 < m_3$;
2. Inverted hierarchy: $m_3 < m_1 < m_2$.

There is another bound on the sum of masses of all light neutrinos:

$$\sum_i m_{\nu_i} < 0.68 \text{ eV} \tag{332}$$

at 95% confidence limit. This comes from cosmology; the analysis of the data collected by the Wilkinson Microwave Anisotropy Probe (WMAP) as well as data from supernovae and galaxy clustering.

8.4 The see-saw mechanism

Neutrino masses are very small, much smaller than even the charged leptons. However, one can always write a Dirac mass term, introducing right-handed neutrinos ν_R and very small Yukawa couplings:

$$\mathcal{L}_D = -m_D \bar{\nu}_L \nu_R + \text{h.c.} \quad \text{with} \;\; m_D = \frac{h_\nu v}{\sqrt{2}} \, . \tag{333}$$

Such neutrinos are called Dirac neutrinos. They conserve the lepton number L: L($\bar{\nu}$) = −L(ν), and ν and $\bar{\nu}$ are different.

One can write another mass term with the ν_R fields:

$$\mathcal{L}_M = -m_R \overline{\nu_R^c} \nu_R + \text{h.c.} \tag{334}$$

Such a term cannot be written for any other SM fermions as they will violate the conservation of electric charge. Both ν_R and $\overline{\nu_R^c}$ annihilate a neutrino and creates an antineutrino, so this mixes ν and $\bar{\nu}$ and violates L. This term is called a Majorana mass term; Majorana fermions are their own antiparticles.

Consider only the electron neutrinos. The observed neutrinos are all light and left-handed, but suppose there is a ν_R field too. These ν_Rs are massive and can have a large Majorana mass. Once the Majorana mass term is present, L is no longer conserved and neutrinos are Majorana particles in nature. The neutrino mass matrix, written in terms of the SU(2)$_L$ doublet and singlet fields, is

$$\begin{pmatrix} \bar{\nu}_L & \bar{\nu}_R \end{pmatrix} \begin{pmatrix} 0 & m_D \\ m_D & m_R \end{pmatrix} \begin{pmatrix} \nu_L \\ \nu_R \end{pmatrix} . \tag{335}$$

Upon diagonalization, the doublet neutrinos get a small mass of m_D^2/M_R (if $m_R \gg m_D$) while the ν_Rs still have almost the same mass m_R. This is known as the *see-saw mechanism* for the induction of neutrino mass.

How do we know if the neutrinos are Dirac or Majorana particles? There is a process called double β decay, where a nucleus emits two electrons and two antineutrinos:

$$X(Z, A) \to Y(Z+2, A) + 2e^- + 2\bar{\nu}_e \, . \tag{336}$$

If neutrinos are their own antiparticles, one of the emitted antineutrinos can be treated as a neutrino and can be absorbed back by another parent nucleus (remember that absorption of particle is the same as emission of antiparticle), so that no neutrinos are emitted:

$$X(Z, A) \to Y(Z+2, A) + 2e^- \, . \tag{337}$$

This is known as a neutrinoless double β decay ($0\nu\beta\beta$) and is a sure sign of Majorana neutrinos. There are a few experiments in progress to check the nature of neutrinos but we have nothing conclusive yet.

While this is outside the scope of this course, if the neutrinos are Majorana particles, the PMNS matrix contains two extra phases known as Majorana phases that cannot be rotated away. However, these phases do not affect the neutrino oscillations.

9 The Structure of Proton

To know what is there inside the proton, we have to probe the proton with photons whose wavelength is comparable to the dimension of the proton. Consider an electron-proton scattering; if the electron is sufficiently energetic, the exchanged photon can probe the proton. There can be two types of such scattering: (i) elastic, $ep \to ep$, and (ii) inelastic, $ep \to eX$, where the proton breaks into fragments.

If the proton were a point particle, the amplitude squared for the elastic scattering would have been the same as of $e\mu \to e\mu$. So let us first digress a bit and recapitulate the QED result.

9.1 Electron-muon scattering in the lab frame

Consider the scattering

$$e(k) + \mu(p) \to e(k') + \mu(p') \tag{338}$$

in the frame where the initial muon was at rest. We will treat the electron to be ultra-relativistic and neglect its mass, but keep that of the muon.

The four-momentum assignments are

$$k = (E, \mathbf{k})\,, \quad k' = (E', \mathbf{k}')\,, \quad p = (M, \mathbf{0})\,, \quad p' = (p_0', \mathbf{p}')\,, \tag{339}$$

and $q = (\nu, \mathbf{q})$ is the four-momentum of the exchanged photon. Let the scattering angle be θ, so that $\mathbf{k}.\mathbf{k}' = |\mathbf{k}||\mathbf{k}'|\cos\theta$. We put $k^2 = k'^2 = 0$ so that $q^2 = (k-k')^2 = -2k.k'$. From four-momentum conservation on the muon vertex, $q + p = p'$, so

$$q^2 = -2p.q = -2\nu M \Rightarrow \nu = E - E' = -\frac{q^2}{2M}\,. \tag{340}$$

We will use the QED result for unpolarized scattering, substitute $p' = k + p - k'$, and also use the relationships above, to get

$$\begin{aligned}
|\mathcal{M}|^2 &= \frac{8e^4}{q^4}\left[(k.p)(k'.p') + (k.p')(k'.p) - M^2(k.k')\right] \\
&= \frac{8e^4}{q^4}\left[(k.p)(k'.k + k'.p) + (k'.p)(k.p - k.k') + \frac{1}{2}M^2q^2\right] \\
&= \frac{8e^4}{q^4}\left[2(k.p)(k'.p) - \frac{1}{2}q^2(k.p - k'.p) + \frac{1}{2}M^2q^2\right] \\
&= \frac{8e^4}{q^4}\left[2EE'M^2 - \frac{1}{2}q^2M(E-E') + \frac{1}{2}M^2q^2\right] \\
&= \frac{8e^4}{q^4}2EE'M^2\left[1 + \frac{q^2}{4EE'} - \frac{q^2}{2M^2}\frac{M(E-E')}{2EE'}\right]\,.
\end{aligned} \tag{341}$$

But $q^2 = 2k.k' = -2EE'(1-\cos\theta) = -4EE'\sin^2(\theta/2)$, and $E - E' = -q^2/2M$, so

$$|\mathcal{M}|^2 = \frac{8e^4}{q^4} 2EE'M^2 \left(\cos^2\frac{\theta}{2} - \frac{q^2}{2M^2}\sin^2\frac{\theta}{2}\right) . \tag{342}$$

It is left as an exercise (not very trivial) to show that

$$\left(\frac{d\sigma}{d\Omega}\right)_{lab} = \left(\frac{\alpha^2}{4E^2\sin^4(\theta/2)}\right)\frac{E'}{E}\left(\cos^2\frac{\theta}{2} - \frac{q^2}{2M^2}\sin^2\frac{\theta}{2}\right) . \tag{343}$$

For small θ, this gives the Rutherford $\sin^{-4}(\theta/2)$ dependence.

Electron is a point particle, so the electron current will retain its structure, but proton is not, so the proton current will no longer have the Lorentz structure of γ^μ. The only four-vectors that can be written are γ^μ and $i\sigma^{\mu\nu}q_\nu$, and their coefficients can be functions of the only scalar available at the proton vertex, q^2. Current conservation forbids $(p-p')^\mu$ as a possible choice, and $(p+p')^\mu$ can be expressed in terms of γ^μ and $i\sigma^{\mu\nu}q_\nu$ (this is known as Gordon decomposition). Thus, we can write

$$j^\mu_{proton} = e\bar{u}(p')\left[F_1(q^2)\gamma^\mu + \frac{\kappa}{2M}F_2(q^2) i\sigma^{\mu\nu}q_\nu\right]u(p)e^{i(p'-p).x} . \tag{344}$$

For long-wavelength photons $q^2 \to 0$, we see only a pointlike proton, so

$$F_1(0) = F_2(0) = 1 \tag{345}$$

and $\kappa = 1.79$ is the anomalous magnetic moment. For neutron, $F_1(0) = 0$, $F_2(0) = 1$, and $\kappa = -1.91$.

Putting the expression for proton current, the differential cross-section turns out to be

$$\left(\frac{d\sigma}{d\Omega}\right)_{lab} = \left(\frac{\alpha^2}{4E^2\sin^4(\theta/2)}\right)\frac{E'}{E}\left\{\left(F_1^2 - \frac{\kappa^2 q^2 F_2^2}{4M^2}\right)\cos^2\frac{\theta}{2} - (F_1+\kappa F_2)^2\frac{q^2}{2M^2}\sin^2\frac{\theta}{2}\right\} . \tag{346}$$

This is known as the Rosenbluth formula. The form factors F_1 and F_2 is nothing but a way to hide our ignorance about the proton structure.

One can also use linear combinations of F_1 and F_2,

$$G_E = F_1 + \frac{\kappa q^2}{4M^2}F_2 , \quad G_M = F_1 + \kappa F_2 , \tag{347}$$

so that

$$\left(\frac{d\sigma}{d\Omega}\right)_{lab} = \left(\frac{\alpha^2}{4E^2\sin^4(\theta/2)}\right)\frac{E'}{E}\left\{\frac{G_E^2 + \tau G_M^2}{1+\tau}\cos^2\frac{\theta}{2} + 2\tau G_M^2\sin^2\frac{\theta}{2}\right\} , \tag{348}$$

where $\tau = -q^2/4M^2$.

Q. Derive (343).

Q. Show that the factor of E'/E in (343) is given by

$$\frac{E'}{E} = \frac{1}{1 + \frac{2E}{M}\sin^2\frac{\theta}{2}} . \tag{349}$$

This is unity if $M \to \infty$ and gives the recoil of the target.

Q. Prove the Gordon decomposition for a fermion of mass m:

$$\bar{u}_f\gamma^\mu u_i = \frac{1}{2m}\bar{u}_f\left[(p_f+p_i)^\mu + i\sigma^{\mu\nu}(p_f-p_i)_\nu\right]u_i . \tag{350}$$

Q. Deduce (348) from (346).

9.2 Inelastic ep scattering

In the process $ep \to eX$, the proton breaks up, so we cannot use the simple form of the proton current. The momentum transfer q^2 is large, so the proton is highly virtual (a real photon has $q^2 = 0$), that is why the process is also called *deep inelastic scattering* (DIS).

The differential cross-section is proportional to the matrix element squared, which in turn contains the product $L_{\mu\nu}H^{\mu\nu}$, where $L_{\mu\nu}$ is the lepton (*i.e.*, electron) tensor (this remains the same as that in $e\mu \to e\mu$ scattering). $H^{\mu\nu}$ is the hadron tensor, which cannot be written like that of elastic scattering, because the proton breaks up and it is not possible to write a single $\bar{u}_f$ for the final-state wavefunction. Rather, we will write the most general form, constructed out of $\eta^{\mu\nu}$ and the independent momenta p (that of the initial proton) and q (the momentum transfer). The cross-section is already summed and averaged over spin, so there are no explicit Dirac matrices. We write

$$H^{\mu\nu} = -W_1\eta^{\mu\nu} + W_2\frac{p^\mu p^\nu}{M^2} + W_4\frac{q^\mu q^\nu}{M^2} + W_5\frac{1}{M^2}(p^\mu q^\nu + q^\mu p^\nu)\,, \tag{351}$$

where the antisymmetric terms are omitted because they vanish when contracted with the symmetric lepton tensor. The W_is are known as *structure functions.* The parity-violating structure function W_3 comes in when we have a neutrino beam and the weak gauge bosons are exchanged, so that has been left out here.

It can be shown that the current is conserved at the hadronic vertex too even though the proton breaks up, so $q_\mu H^{\mu\nu} = q_\nu H^{\mu\nu} = 0$. This gives

$$q_\mu H^{\mu\nu} = 0 = \left[-W_1 + \frac{W_4}{M^2}q^2 + \frac{W_5}{M^2}p.q\right]q^\nu + \left[\frac{W_2}{M^2}p.q + \frac{W_5}{M^2}q^2\right]p^\nu\,. \tag{352}$$

Equating the coefficients of q^ν and p^ν separately to zero,

$$W_5 = -\frac{p.q}{q^2}W_2\,,\quad W_4 = \left(\frac{p.q}{q^2}\right)^2 W_2 + \frac{M^2}{q^2}W_1\,. \tag{353}$$

Thus, one can write the hadronic tensor as

$$H^{\mu\nu} = W_1\left(-\eta^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) + \frac{W_2}{M^2}\left(p^\mu - \frac{p.q}{q^2}q^\mu\right)\left(p^\nu - \frac{p.q}{q^2}q^\nu\right)\,. \tag{354}$$

The W_is are functions of independent Lorentz scalars, but now there are two, q^2 and $p.q$. The second variable can also be substituted by $\nu = p.q/M$. The invariant mass of the fragments, W, is related to q^2 and M by

$$W^2 = (p+q)^2 = M^2 + 2M\nu + q^2\,. \tag{355}$$

One generally uses two other variables x and y in place of q^2 and $p.q$, defined as

$$x = -q^2/2p.q = -q^2/2M\nu\,,\quad y = (p.q)/(p.k)\,. \tag{356}$$

$q^2 < 0$ and $\nu \geq 0$, so $x \geq 0$. Also, $W^2 \geq M^2$, so $x \leq 1$, where $x = 1$ corresponds to elastic scattering $W^2 = M^2$. y is a Lorentz invariant quantity, so we can use any frame to evaluate it. In the lab frame where the proton is at rest, $y = (E - E')/E$, but $E' \leq E$, so $0 \leq y \leq 1$.

Using the standard form of $L_{\mu\nu}$, we get

$$L_{\mu\nu}H^{\mu\nu} = 4W_1(k.k') + \frac{2W_2}{M^2}\left[2(p.k)(p.k') - M^2 k.k'\right]\,, \tag{357}$$

where W_1 and W_2 are functions of ν and q^2. In the lab frame, the above equation becomes

$$L_{\mu\nu}H^{\mu\nu} = 4EE'\left[W_2\cos^2(\theta/2) + 2W_1\sin^2(\theta/2)\right] . \tag{358}$$

Inserting the flux factors and the phase space, one has

$$\begin{aligned}\left(\frac{d\sigma}{d\Omega\, dE'}\right)_{lab} &= \left(\frac{\alpha^2}{4E^2\sin^4(\theta/2)}\right)\left\{W_2(\nu,q^2)\cos^2\frac{\theta}{2} + 2W_1(\nu,q^2)\sin^2\frac{\theta}{2}\right\} \\ &= \frac{\alpha^2}{q^4}\frac{E'}{E}L_{\mu\nu}H^{\mu\nu} ,\end{aligned} \tag{359}$$

where we have used $q^2 = (k-k')^2 = -2EE'(1-\cos\theta)$ and (358).

Let us then take stock of what we have found.

For point particle scattering like $e\mu \to e\mu$,

$$\frac{d\sigma}{dE'\, d\Omega} = \frac{4\alpha^2 E'^2}{q^4}\mathcal{P} \tag{360}$$

where

$$\mathcal{P} = \left(\cos^2\frac{\theta}{2} - \frac{q^2}{2m^2}\sin^2\frac{\theta}{2}\right)\delta\left(\nu + \frac{q^2}{2m}\right) . \tag{361}$$

The δ-function tells you that this is a $2 \to 2$ scattering and the final-state muon momentum p' is related to the initial-state muon momentum p by $p' = p + q$. The integration over E' is taken care of by the δ-function and we recover the familiar expressions of $d\sigma/d\Omega$.

For elastic scattering $ep \to ep$, we have an identical expression, with

$$\mathcal{P} = \left(\frac{G_E^2 + \tau G_M^2}{1+\tau}\cos^2\frac{\theta}{2} + 2\tau G_M^2\sin^2\frac{\theta}{2}\right)\delta\left(\nu + \frac{q^2}{2m}\right) . \tag{362}$$

For inelastic scattering, surprisingly, we still have the same form, but now there is no momentum-conserving δ-function:

$$\mathcal{P} = \left(W_2(\nu,q^2)\cos^2\frac{\theta}{2} + 2W_1(\nu,q^2)\sin^2\frac{\theta}{2}\right) . \tag{363}$$

Thus, for a pointlike proton — the thing that a long-wavelength photon will see — the structure functions are

$$2W_1 = -\frac{q^2}{2M^2}\delta\left(\nu + \frac{q^2}{2M}\right) , \quad W_2 = \delta\left(\nu + \frac{q^2}{2M}\right) . \tag{364}$$

10 The Parton Model

There is an intriguing property of (364). Using the identity $\delta(x/a) = a\delta(x)$, and substituting $Q^2 = -q^2$ (so that $Q^2 > 0$), we can rewrite (364) as

$$2MW_1(\nu,Q^2) = \frac{Q^2}{2M\nu}\delta\left(1 - \frac{Q^2}{2M\nu}\right) = x\delta(1-x) , \quad \nu W_2(\nu,Q^2) = \delta\left(1 - \frac{Q^2}{2M\nu}\right) = \delta(1-x) . \tag{365}$$

Note an interesting property: these point-like structure functions are functions of the dimensionless variable $x \equiv Q^2/2M\nu$ only (see (356)), and not the individual variables Q^2 and ν. In general the elastic scattering form factors depend on Q^2, which is a sign that proton is not a point particle. However, in the high-Q^2 limit when the proton breaks, the structure functions again show point-like behaviour, they depend only on x but not on the precise value of Q^2. So in the limit $Q^2, \nu \to \infty$ keeping x fixed,

$$W_i(x, Q^2) \longrightarrow W_i(x) \, . \tag{366}$$

This is known as *Bjorken scaling* and is an indication that there are point-like constituents inside the proton with which the virtual photon is interacting now. These point like constituents are called *partons*; they include the valence quarks, the sea quarks and antiquarks (see later), and the gluons too. Of course, scaling is only a leading-order phenomenon, it is violated by higher-order effects of QCD, in particular by gluon emission. We will come back to that soon.

Before we move any further, let us see what happens if the photon strikes a pointlike parton whose mass is $m = zM$, where $0 < z \leq 1$. All Ms in (365) will be replaced by zM, so that the argument of the δ-function will be $1 - x/z$. Thus, $x = z$, or in other words, if $Q^2/2M\nu = x$, only that parton with mass xM will be excited, so this is like an elastic scattering with the parton.

Thus, for high Q^2, the photon sees the partons, and not the proton as a whole. In this limit, the structure functions can be written as form factors (not to be confused with the elastic $ep \to ep$ form factors, although this is an elastic scattering):

$$MW_1(\nu, Q^2) \to F_1(x) \, , \quad \nu W_2(\nu, Q^2) \to F_2(x) \, , \tag{367}$$

so that, from (365),

$$F_2(x) = 2x \, F_1(x) \, . \tag{368}$$

This is known as the *Callan-Gross relation.*

10.1 Parton kinematics

Consider a parton of mass m carrying a fraction x of the proton momentum $\mathbf{p}$. Even this simple sentence needs clarification as $\mathbf{p}$ is a vector. Let the energy of the initial proton be E, and its longitudinal momentum be p_L. Obviously, there is no momentum perpendicular to its direction of motion, so p_T, the transverse momentum, is identically zero, and $p_L = |\mathbf{p}|$. From our earlier discussion, the photon that hit the parton must have $Q^2/2M\nu = x$.

We now go into an inertial frame where $|\mathbf{p}| \gg m, M$. This is known as the *infinite momentum frame.* In this frame, all decay products of the proton are highly boosted in the forward direction, so the emitted parton will have longitudinal momentum xp_L and transverse momentum zero [22]. As we can neglect all masses, the energy of the parton is $xp_L = xE$.

If $f_i(x)$ is the probability that the struck parton i carries a momentum fraction x, then

$$\sum_i \int x f_i(x) \, dx = 1 \tag{369}$$

---

[22] In the lab frame, for example, the parton must have some nonzero p_T. The decay products must have some angle between them to conserve both energy and momentum. However, one can always boost back from the infinite momentum frame to the lab frame.

as the momentum fractions must sum up to one. This sum runs over all partons, both charged (quarks and antiquarks) and neutral (gluons). The electron-proton scattering can be thought schematically as

$$\sum_i \int dx\, e_i^2 \times \text{photon} - \text{parton scattering} \tag{370}$$

where e_i is the charge of the i-th parton; this factor comes because the interaction is through photons. Obviously, this takes out the neutral partons.

Now let us use (365) for photon-parton scattering, where the parton carries a momentum fraction z. For a single parton,

$$\nu W_2 \to F_2 = \delta\left(1 - \frac{x}{z}\right) = z\delta(z - x)\,, \tag{371}$$

and for the entire proton

$$F_2(x) = \sum_i \int dz\, e_i^2 f_i(z) z\delta(z - x) = \sum_i e_i^2 x f_i(x)\,, \qquad F_1(x) = \frac{1}{2x} F_2(x)\,. \tag{372}$$

The Callan-Gross relation is a direct consequence of the fact that the quarks are spin-$\frac{1}{2}$ objects. Consider a head-on collision with a photon which reverses the quark momentum. Helicity is conserved in high-energy interactions, so the spin direction also flips. This can only happen if the photon has helicity $\lambda = \pm 1$, *i.e.* transverse. Photons with helicity $\lambda = 0$ [23] can never flip the spin direction, so we expect the cross-section with longitudinal photons to vanish in the limit $\nu, Q^2 \to \infty$.

Q. (*) For virtual photons, one can write the polarization vectors as

$$\begin{aligned} \lambda = \pm 1 \quad &\Rightarrow \qquad \epsilon_\pm = \mp \frac{1}{\sqrt{2}}(0, 1, \pm i, 0)\,, \\ \lambda = 0 \quad &\Rightarrow \qquad \epsilon_0 = \frac{1}{\sqrt{-q^2}}\left(\sqrt{\nu^2 - q^2}, 0, 0, \nu\right)\,. \end{aligned} \tag{373}$$

Show that $q.\epsilon = 0$ for each λ, and for a spacelike photon ($q^2 < 0$)

$$\sum_{\lambda = 0, \pm 1} (-1)^{\lambda + 1} \epsilon^{\mu *} \epsilon^\nu = -\eta^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\,. \tag{374}$$

Now use the form of (354) to show that

$$\frac{\sigma_L}{\sigma_T} = \frac{\frac{1}{2}(\sigma_+ + \sigma_-)}{\sigma_0} = \frac{\left(1 - \frac{\nu^2}{q^2}\right) W_2 - W_1}{W_1}\,. \tag{375}$$

Evaluate in the lab frame where $p = (M, \mathbf{0})$ and $q = (\nu, 0, 0, \sqrt{\nu^2 - q^2})$.

Q. In the limit $\nu, Q^2 \equiv -q^2 \to \infty$ with x fixed, show that one can neglect 1 compared to ν^2/q^2 in the numerator of (375). Hence show that we can write (375) for pointlike partons as

$$\frac{\sigma_L}{\sigma_T} = \frac{\nu^2 W_2 - Q^2 W_1}{Q^2 W_1} = \frac{F_2}{2x F_1} - 1 \to 0\,. \tag{376}$$

---

[23] These are virtual photons, so there can be a longitudinal component too.

10.2 Parton distribution functions

Let us consider electron-proton scattering again, and use (372). If there were only three quarks inside the proton and nothing else, we would have found

$$\frac{1}{x}F_2^{ep}(x) = \left(\frac{2}{3}\right)^2 u^p(x) + \left(\frac{1}{3}\right)^2 d^p(x)\,, \tag{377}$$

where $q^p(x)$ is the probability distribution function f for the quark q inside a proton. There are two up quarks and one down quark, so

$$\int_0^1 dx\, u^p(x) = 2\,, \quad \int_0^1 dx\, d^p(x) = 1\,. \tag{378}$$

In fact, if there are no gluons, the quarks cannot "talk" with each other and exchange momentum, so all of them will be stuck with $x = \frac{1}{3}$; $F_2^{ep}(x)$ will show a δ-function spike at $x = \frac{1}{3}$.

The situation is more complicated. There are gluons inside the proton, also quark-antiquark pairs created from the gluons. For the time being, let us neglect the possibility that the proton can contain anything heavier than the strange quark. We can now write

$$\frac{1}{x}F_2^{ep}(x) = \left(\frac{2}{3}\right)^2 \left[u^p(x) + \bar{u}^p(x)\right] + \left(\frac{1}{3}\right)^2 \left[d^p(x) + \bar{d}^p(x)\right] + \left(\frac{1}{3}\right)^2 \left[s^p(x) + \bar{s}^p(x)\right]\,. \tag{379}$$

These six functions are known as parton distribution functions (PDF). The PDF for gluons do not come into the picture as they do not have any electric charge. One can get a similar expression for neutrons by scattering electrons off deuterium:

$$\frac{1}{x}F_2^{en}(x) = \left(\frac{2}{3}\right)^2 \left[u^n(x) + \bar{u}^n(x)\right] + \left(\frac{1}{3}\right)^2 \left[d^n(x) + \bar{d}^n(x)\right] + \left(\frac{1}{3}\right)^2 \left[s^n(x) + \bar{s}^n(x)\right]\,. \tag{380}$$

But there is an isospin symmetry between proton and neutron, so

$$u^p(x) = d^n(x) \equiv u(x)\,, \quad d^p(x) = u^n(x) \equiv d(x)\,, \quad s^p(x) = s^n(x) \equiv s(x)\,. \tag{381}$$

One can get a bound on the ratio of these two form factors:

$$\frac{1}{4} \le \frac{F_2^{en}(x)}{F_2^{ep}(x)} \le 4 \tag{382}$$

where the lower limit is for the case when there are only d quarks inside the proton, and the upper limit is for when there are only u quarks.

Only $u(x)$ and $d(x)$ will have contributions from valence quarks as well as $q\bar{q}$ pairs created out of gluons; these are called *sea quarks*. For the sea, quark and antiquark contributions should be identical. Also, all flavours should contribute equally [24], so

$$u(x) = u_v(x) + u_s(x)\,, \quad d(x) = d_v(x) + d_s(x)\,, \tag{383}$$

and

$$u_s(x) = \bar{u}_s(x) = d_s(x) = \bar{d}_s(x) = s_s(x) = \bar{s}_s(x) \equiv S(x)\,. \tag{384}$$

---

[24] They do not. In particular, s is heavier than u or d, so there will be less number of $s\bar{s}$ pairs. But this is only a leading approximation.

The normalization now becomes

$$\int_0^1 dx\ [u(x) - \bar{u}(x)] = 2\,, \quad \int_0^1 dx\ [d(x) - \bar{d}(x)] = 1\,, \quad \int_0^1 dx\ [s(x) - \bar{s}(x)] = 0\,. \tag{385}$$

But $u(x) - \bar{u}(x) = u_v(x)$, and $d(x) - \bar{d}(x) = d_v(x)$; this is what we mean when we say that the proton contains two up quarks and one down quark.

From (379) and (380), we get

$$\frac{1}{x}F_2^{ep}(x) = \frac{1}{9}(4u_v + d_v) + \frac{4}{3}S\,, \quad \frac{1}{x}F_2^{en}(x) = \frac{1}{9}(u_v + 4d_v) + \frac{4}{3}S\,. \tag{386}$$

At low x, the number of gluons increase; this is completely analogous to the excess of soft photons in bremsstrahlung. If you remember, in a simplistic treatment, the number of bremsstrahlung photons blow up as the energy goes to zero; no extra energy is needed to create an infinite number of zero-energy photons. Similarly, as $x \to 0$, the gluon distribution function starts growing. As the sea quarks are created out of gluons, we expect $S(x)$ to grow at low x and swamp the valence quark contributions, so as $x \to 0$, $F_2^{ep} = F_2^{en}$. On the other hand, for large momentum transfer ($x \to 1$), the valence quarks dominate. Both these facts are borne out by experiment.

One can also construct an observable without the sea quark contribution:

$$\frac{1}{x}\left(F_2^{ep}(x) - F_2^{en}(x)\right) = \frac{1}{3}(u_v - d_v)\,. \tag{387}$$

The total momentum of the proton must be carried by all the partons. Summing up the momentum of the quark-antiquark partons,

$$\int_0^1 dx\,(xp)\left[u + \bar{u} + d + \bar{d} + s + \bar{s}\right] = p - p_g \tag{388}$$

where p_g is the momentum carried by the gluons. Dividing by p,

$$\int_0^1 dx\,x\left[u + \bar{u} + d + \bar{d} + s + \bar{s}\right] = 1 - \epsilon_g \tag{389}$$

where $\epsilon_g = p_g/p$ is the momentum fraction carried by the gluons. From the experiments, we have

$$\int_0^1 dx\,F_2^{ep}(x) = \frac{4}{9}\epsilon_u + \frac{1}{9}\epsilon_d + \frac{1}{9}\epsilon_s = 0.18\,, \quad \int_0^1 dx\,F_2^{en}(x) = \frac{1}{9}\epsilon_u + \frac{4}{9}\epsilon_d + \frac{1}{9}\epsilon_s = 0.12\,, \tag{390}$$

with

$$\epsilon_u = \int_0^1 dx\,x(u + \bar{u}) \tag{391}$$

is the momentum fraction carried by the up quarks (both valence and sea), and so on. The strange quarks occur only in the sea, and their contributions to the total momentum of the proton is small. Neglecting ϵ_s, we get

$$\epsilon_u = 0.36\,, \quad \epsilon_d = 0.18\,, \quad \epsilon_g = 0.46\,. \tag{392}$$

Thus, almost half of the proton momentum is carried by the gluons. For example, at the LHC running at $\sqrt{s} = 8$ TeV, each proton has an energy of 4 TeV, so about 2 TeV of it is carried by the gluons. Thus, the process $gg \to X$ can create some new particle X only with a mass less than about 4 TeV. For pair production, this drops to 2 TeV. The actual reach is even lower as the gluon distribution function falls off at large x.

The gluon PDF is not probed by the photons; this is estimated by how the gluons are radiated off quarks and other gluons. This is an extremely important topic for hadron colliders but is unfortunately outside the scope of the course.